

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 23:33:41 ; Search time 27 Seconds
(without alignments)
2192.437 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877
Sequence: 1 NTVRLKLYSLRWISDHEYL.....AQHQHYTHMGHFIKQCFSLP 716

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	755	4	PCT-US93-07923-3
2	3877	100.0	759	4	PCT-US93-07923-2
3	3877	100.0	766	1	US-08-230-491A-3
4	3877	100.0	766	1	US-08-619-280A-3
5	3877	100.0	766	1	US-08-940-391-3
6	3877	100.0	766	2	US-09-794-236-1
7	3877	100.0	766	2	US-09-265-606-3
8	3877	100.0	775	2	US-09-949-016-10450
9	3871	99.8	766	2	US-10-002-593-6
10	3871	99.8	766	2	US-09-949-016-6146
11	3871	99.8	766	2	US-10-423-714-6
12	3867	99.7	766	2	US-09-518-550-27
13	2948	76.0	593	4	PCT-US93-07923-11
14	2138	55.1	760	1	US-08-230-491A-2
15	2138	55.1	760	1	US-08-619-280A-2
16	2138	55.1	760	1	US-08-940-391-2
17	2138	55.1	760	2	US-09-265-606-2
18	2136.5	55.1	761	2	US-09-518-550-26
19	1229.5	31.7	547	2	US-09-949-016-8330
20	1217	31.4	786	2	US-09-976-674-5
21	1198	30.9	743	2	US-10-363-937-4
22	1196	30.8	706	2	US-09-976-674-41
23	1152.5	29.7	691	2	US-09-976-674-43
24	987	25.5	771	2	US-09-079-592-2
25	963.5	24.9	771	2	US-09-462-284-2
26	908	23.4	818	2	US-09-462-845-3
27	908	23.4	818	2	US-10-402-312-3

Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 42, Appli
Sequence 29, Appli
Sequence 28, Appli
Sequence 11, Appli
Sequence 45296, A
Sequence 19472, A
Sequence 1, Appli
Sequence 1, Appli
Sequence 19924, A
Sequence 30, Appli
Sequence 33, Appli
Sequence 35, Appli
Sequence 3, Appli
Sequence 23, Appli
Sequence 27, Appli

ALIGNMENTS

RESULT 1
PCT-US93-07923-3
; Sequence 3, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlossman, Stuart F.
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 755
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-07923-3

Query Match 100.0%; Score 3877; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEFHGSINDYSISPD 60
Db 40 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEFHGSINDYSISPD 99
QY 61 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 120
Db 100 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 159
QY 121 DIYVKIEPNLPSRITWTGKEDIYNGITDWYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 160 DIYVKIEPNLPSRITWTGKEDIYNGITDWYEEVFSAYSALWSPNGTFLAYAQFND 219
QY 181 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSVVTNATSIQITAP 240
Db 220 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSVVTNATSIQITAP 279
QY 241 ASMLIGDHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWNCILVARQHIE MST 300
Db 280 ASMLIGDHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWNCILVARQHIE MST 339
QY 301 TGVWGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 360
Db 340 TGVWGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 399
QY 361 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
Db 400 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 459
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEONSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 480
Db 460 RCGSGPLPLYTLHSSVNDKGLRVLEONSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 519
QY 481 PPHFDKSKKYPILLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDRGSGYQGDKI 540
Db 520 PPHFDKSKKYPILLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDRGSGYQGDKI 579
QY 541 MHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSMLGSGGVFKCG 600
Db 580 MHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSMLGSGGVFKCG 639
QY 601 IAVAPVSRWEYDYSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 640 IAVAPVSRWEYDYSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 699
QY 661 VHFQOQAISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHFIIKQCFSLP 716
Db 700 VHFQOQAISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHFIIKQCFSLP 755
```

RESULT 2

```
PCT-US93-07923-2
; Sequence 2, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlossman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; OPERATING SYSTEM: IBM PS/2 Model 502 or 55SX
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
```

```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janle K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 759
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; PCT-US93-07923-2

Query Match 100.0%; Score 3877; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEFHGSINDYSISPD 60
Db 44 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDEFHGSINDYSISPD 103
QY 61 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 120
Db 104 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 163
QY 121 DIYVKIEPNLPSRITWTGKEDIYNGITDWYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 164 DIYVKIEPNLPSRITWTGKEDIYNGITDWYEEVFSAYSALWSPNGTFLAYAQFND 223
QY 181 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSVVTNATSIQITAP 240
Db 224 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSVVTNATSIQITAP 283
QY 241 ASMLIGDHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWNCILVARQHIE MST 300
Db 284 ASMLIGDHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWNCILVARQHIE MST 343
QY 301 TGVWGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 360
Db 344 TGVWGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 403
QY 361 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
Db 404 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 463
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEONSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 480
Db 464 RCGSGPLPLYTLHSSVNDKGLRVLEONSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 523
QY 481 PPHFDKSKKYPILLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDRGSGYQGDKI 540
Db 524 PPHFDKSKKYPILLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDRGSGYQGDKI 583
QY 541 MHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSMLGSGGVFKCG 600
Db 584 MHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSMLGSGGVFKCG 643
QY 601 IAVAPVSRWEYDYSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 644 IAVAPVSRWEYDYSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 703
QY 661 VHFQOQAISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHFIIKQCFSLP 716
Db 704 VHFQOQAISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHFIIKQCFSLP 759
```

```
RESULT 3
US-08-230-491A-3
; Sequence 3, Application US/08230491A
; Patent No. 5587299
; GENERAL INFORMATION:
; APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;
; APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR
; TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FELFE & LYNCH
; STREET: 805 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT - ASC II
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/230,491A
; FILING DATE: 20-APRIL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5587299man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 330
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-230-491A-3
Query Match 100.0%; Score 3877; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 60
Db 51 NTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 110
Qy 61 GQFTLLVYVYKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVWSPVGHKLAYVWNN 120
Db 111 GQFILLVYVYKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTYEEVFSAYSLAWSPNGTFLAYAGFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTYEEVFSAYSLAWSPNGTFLAYAGFND 230
Qy 181 TEVPLIEYSFYSDSLQPKTVRPYPKAGAVNPTKFFVNTDLSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDSLQPKTVRPYPKAGAVNPTKFFVNTDLSLSSVTNATSIQITAP 290
Qy 241 ASMLIGHLYCDVTWATQERISLOWLRRIQNYSDICDYDESSGRWNCVLAQRIEMST 300
Db 291 ASMLIGHLYCDVTWATQERISLOWLRRIQNYSDICDYDESSGRWNCVLAQRIEMST 350
Qy 301 TGWVGRFRPSEPHTLDGNSFYKIIISNEBGRHICVFQIDKDKCTFTTKGTWEVIGEAL 360
Db 351 TGWVGRFRPSEPHTLDGNSFYKIIISNEBGRHICVFQIDKDKCTFTTKGTWEVIGEAL 410
Qy 361 TSDLYLYISNEYKGMPCGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSPSKEAKYYQL 420
Db 51 NTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 110

Db 411 TSDLYLYISNEYKGMPCGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSPSKEAKYYQL 470
Qy 421 RCSPGLPLTYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDLFIILNETKFWYQML 480
Db 471 RCSPGLPLTYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDLFIILNETKFWYQML 530
Qy 481 PPFPDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDGRSGSYQSGDKI 540
Db 531 PPFPDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDGRSGSYQSGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSNVLGSGSGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSNVLGSGSGVFKCG 650
Qy 601 IAVAPVSRWEYSDVYTERYMGTPEDNLDHYRNSVTMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYSDVYTERYMGTPEDNLDHYRNSVTMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
Db 711 VHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 4
US-08-619-280A-3
; Sequence 3, Application US/08619280A
; Patent No. 5767242
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280A
; FILING DATE: 18-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5767242man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-619-280A-3
Query Match 100.0%; Score 3877; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 60
Db 51 NTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 110
```

QY 61 GQFILLVYVQWRHSYTSYSDIYDLNKRQLITEERI PNTQWTVSPVGHKLA YVWNN 120
DB 111 GQFILLVYVQWRHSYTSYSDIYDLNKRQLITEERI PNTQWTVSPVGHKLA YVWNN 170
QY 121 DIYVKIEPNLPSYRI TWGKEDI IYNGITDWYEEVEFSAYSALWSPNGTFLAYAQFND 180
DB 171 DIYVKIEPNLPSYRI TWGKEDI IYNGITDWYEEVEFSAYSALWSPNGTFLAYAQFND 230
QY 181 TEVPLIEYSFYSDLSQPKTVRPVYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 240
DB 231 TEVPLIEYSFYSDLSQPKTVRPVYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 290
QY 241 ASMLIGHYICDVWTATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIE MST 300
DB 291 ASMLIGHYICDVWTATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIE MST 350
QY 301 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFITKGTWEVIGIEAL 360
DB 351 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFITKGTWEVIGIEAL 410
QY 361 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYIQL 420
DB 411 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYIQL 470
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 480
DB 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 530
QY 481 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 540
DB 531 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 590
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGSYGGYVTSMLVSGSGYVFKCG 600
DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGSYGGYVTSMLVSGSGYVFKCG 650
QY 601 IAVAPVSRWEYYSVYTERYMG LPTPEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPVSRWEYYSVYTERYMG LPTPEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQOQAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF1KQCFSLP 716
DB 711 VHFQOQAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF1KQCFSLP 766

RESULT 5

US-08-940-391-3
; Sequence 3, Application US/08940391
; Patent No. 5965373
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION
; TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,391
; FILING DATE: 01-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:

RESULT 6

US-09-794-236-1

; Sequence 1, Application US/09794236
; Patent No. 6337069
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/794,236
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-236-1

Query Match 100.0%; Score 3877; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	NTYRLKLYSLRWISDHELYLKQENNILVFNAYGNSVFLNSTFDFGHSINDYSISPD	60
Db	51	NTYRLKLYSLRWISDHELYLKQENNILVFNAYGNSVFLNSTFDFGHSINDYSISPD	110
Qy	61	GQFILLENNYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN	120
Db	111	GQFILLENNYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN	170
Qy	121	DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSAWSPNGTFLAYAQFND	180
Db	171	DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSAWSPNGTFLAYAQFND	230
Qy	181	TEVPLIEYSFSDLSQYPKTVRYPYKAGAVNPTVKFFVNTDLSSTVNTATSIQTAP	240
Db	231	TEVPLIEYSFSDLSQYPKTVRYPYKAGAVNPTVKFFVNTDLSSTVNTATSIQTAP	290
Qy	241	ASMLIGHYLCVDTWATQERISLOWLRRIQNSYVMDICDYDESSGRWNCVLAHQHIE	300
Db	291	ASMLIGHYLCVDTWATQERISLOWLRRIQNSYVMDICDYDESSGRWNCVLAHQHIE	350
Qy	301	TGWVGRFRPSEPHTFLDGNFSYKIIISNEEGYRHCYFQIDKDKCTFITKGTWEVIGIEAL	360
Db	351	TGWVGRFRPSEPHTFLDGNFSYKIIISNEEGYRHCYFQIDKDKCTFITKGTWEVIGIEAL	410
Qy	361	TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL	420
Db	411	TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL	470
Qy	421	RCSGPGLPLVTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWYQML	480
Db	471	RCSGPGLPLVTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWYQML	530
Qy	481	PPHFDKSKYPLLDVYVAGPCSKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI	540
Db	531	PPHFDKSKYPLLDVYVAGPCSKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI	590
Qy	541	MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTVSNVLGSGGVKPCG	600
Db	591	MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTVSNVLGSGGVKPCG	650
Qy	601	IAPVSRWEYSDSVYTERYMGFLPTPEDNLDHNRNSTVMSRAENFKOVEYLLHGTADDN	660
Db	651	IAPVSRWEYSDSVYTERYMGFLPTPEDNLDHNRNSTVMSRAENFKOVEYLLHGTADDN	710
Qy	661	VHFQSAQISKALVDVGVDFQAMWYTTDEDHGIASTAHQHIYTHMSHFIKQCFSLP	716
Db	711	VHFQSAQISKALVDVGVDFQAMWYTTDEDHGIASTAHQHIYTHMSHFIKQCFSLP	766

RESULT 7
US-09-265-606-3

; Sequence 3, Application US/09265606
; Patent No. 6846910
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6846910man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-3200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-265-606-3

Query Match 100.0%; Score 3877; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	NTYRLKLYSLRWISDHELYLKQENNILVFNAYGNSVFLNSTFDFGHSINDYSISPD	60
Db	51	NTYRLKLYSLRWISDHELYLKQENNILVFNAYGNSVFLNSTFDFGHSINDYSISPD	110
Qy	61	GQFILLENNYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN	120
Db	111	GQFILLENNYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN	170
Qy	121	DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSAWSPNGTFLAYAQFND	180
Db	171	DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSAWSPNGTFLAYAQFND	230
Qy	181	TEVPLIEYSFSDLSQYPKTVRYPYKAGAVNPTVKFFVNTDLSSTVNTATSIQTAP	240
Db	231	TEVPLIEYSFSDLSQYPKTVRYPYKAGAVNPTVKFFVNTDLSSTVNTATSIQTAP	290
Qy	241	ASMLIGHYLCVDTWATQERISLOWLRRIQNSYVMDICDYDESSGRWNCVLAHQHIE	300
Db	291	ASMLIGHYLCVDTWATQERISLOWLRRIQNSYVMDICDYDESSGRWNCVLAHQHIE	350
Qy	301	TGWVGRFRPSEPHTFLDGNFSYKIIISNEEGYRHCYFQIDKDKCTFITKGTWEVIGIEAL	360
Db	351	TGWVGRFRPSEPHTFLDGNFSYKIIISNEEGYRHCYFQIDKDKCTFITKGTWEVIGIEAL	410
Qy	361	TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL	420

Db 411 TSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKLDPIILNETKFWQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKLDPIILNETKFWQMIL 530
Qy 481 PPFPDKSKYPLLLDVVAGPCSKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 540
Db 531 PPFPDKSKYPLLLDVVAGPCSKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 590
Qy 541 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLVSGSGGVFKCG 600
Db 591 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLVSGSGGVFKCG 650
Qy 601 IAVAPVSRWEYDVSYYTERYMGILPTPEDNLDHYRSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSYYTERYMGILPTPEDNLDHYRSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 8
US-09-949-016-10450
; Sequence 10450, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10450
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10450

Query Match 100.0%; Score 3877; DB 2; Length 775;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHELYLKQENNLIVFNAEYGNSSVFLNENSTFDFGHSINDYISPD 60
Db 60 NTYRLKLYSLRWISDHELYLKQENNLIVFNAEYGNSSVFLNENSTFDFGHSINDYISPD 119
Qy 61 GQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 120
Db 120 GQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 179
Qy 121 DIYVKIEPNLPYSRITWTGKEDIYNGITDMWYEEVFSAYLSALWSPNGTFLAYAQFND 180
Db 180 DIYVKIEPNLPYSRITWTGKEDIYNGITDMWYEEVFSAYLSALWSPNGTFLAYAQFND 239
Qy 181 TEVPLIEYSFYSDLSQYKPTVRVPYKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 240
Db 240 TEVPLIEYSFYSDLSQYKPTVRVPYKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 299
Qy 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIE MST 300
Db 300 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIE MST 359
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFITKGTWEVIGIAL 360

Db 360 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFITKGTWEVIGIAL 419
Qy 361 TSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 420 TSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 479
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKLDPIILNETKFWQMIL 480
Db 480 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKLDPIILNETKFWQMIL 539
Qy 481 PPFPDKSKYPLLLDVVAGPCSKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 540
Db 540 PPFPDKSKYPLLLDVVAGPCSKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 599
Qy 541 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLVSGSGGVFKCG 600
Db 600 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLVSGSGGVFKCG 659
Qy 601 IAVAPVSRWEYDVSYYTERYMGILPTPEDNLDHYRSTVMSRAENFKQVEYLLIHGTADDN 660
Db 660 IAVAPVSRWEYDVSYYTERYMGILPTPEDNLDHYRSTVMSRAENFKQVEYLLIHGTADDN 719
Qy 661 VHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
Db 720 VHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 775

RESULT 9
US-10-002-593-6
; Sequence 6, Application US/10002593
; Patent No. 6586198
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING
; INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Acty Docket No. 6586198 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-593-6

Query Match 99.8%; Score 3871; DB 2; Length 766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHELYLKQENNLIVFNAEYGNSSVFLNENSTFDFGHSINDYISPD 60
Db 51 NTYRLKLYSLRWISDHELYLKQENNLIVFNAEYGNSSVFLNENSTFDFGHSINDYISPD 110
Qy 61 GQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 120
Db 111 GQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPYSRITWTGKEDIYNGITDMWYEEVFSAYLSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPYSRITWTGKEDIYNGITDMWYEEVFSAYLSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYKPTVRVPYKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYKPTVRVPYKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 290
Qy 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIE MST 300
Db 291 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIE MST 350

QY 301 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFITKGTWVEIGIEAL 360
DB 351 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFITKGTWVEIGIEAL 410
QY 361 TSDLYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYOL 420
DB 411 TSDLYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYOL 470
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFFILNETKFWYQWIL 480
DB 471 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFFILNETKFWYQWIL 530
QY 481 PPFDKSKKYPPLLDVYAGPCSKADTVFRLNWTYLASTENIIIVASFDFGRGSGYQGDKI 540
DB 531 PPFDKSKKYPPLLDVYAGPCSKADTVFRLNWTYLASTENIIIVASFDFGRGSGYQGDKI 590
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVFKCG 600
DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVFKCG 650
QY 601 IAVAPSRWEYDSVYTERYMGFLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPSRWEYDSVYTERYMGFLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 10
US-09-949-016-6146
; Sequence 6146, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6146
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6146

Query Match 99.8%; Score 3871; DB 2; Length 766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHELYLKQENNLVFNNAEYGNSSVFLNSTFDFEGHSINDYSISPD 60
DB 51 NTYRLKLYSLRWISDHELYLKQENNLVFNNAEYGNSSVFLNSTFDFEGHSINDYSISPD 110
QY 61 GQFILLYNNYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 120
DB 111 GQFILLYNNYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYLSALWSPNGTFLAYAQFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYLSALWSPNGTFLAYAQFND 230
QY 181 TEVPLIEYSFYSDESLOQPKTVRPVPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 240

DB 231 TEVPLIEYSFYSDESLOQPKTVRPVPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 290
QY 241 ASMLIGHYLCDYTWATQERISLOWLRRIONYSVMDICDYDESSGRNWCIVARQHIEMST 300
DB 291 ASMLIGHYLCDYTWATQERISLOWLRRIONYSVMDICDYDESSGRNWCIVARQHIEMST 350
QY 301 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFITKGTWVEIGIEAL 360
DB 351 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFITKGTWVEIGIEAL 410
QY 361 TSDLYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYOL 420
DB 411 TSDLYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYOL 470
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFFILNETKFWYQWIL 480
DB 471 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFFILNETKFWYQWIL 530
QY 481 PPFDKSKKYPPLLDVYAGPCSKADTVFRLNWTYLASTENIIIVASFDFGRGSGYQGDKI 540
DB 531 PPFDKSKKYPPLLDVYAGPCSKADTVFRLNWTYLASTENIIIVASFDFGRGSGYQGDKI 590
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVFKCG 600
DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVFKCG 650
QY 601 IAVAPSRWEYDSVYTERYMGFLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPSRWEYDSVYTERYMGFLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 11
US-10-423-714-6
; Sequence 6, Application US/10423714
; Patent No. 6887679
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERT
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Atty Docket No. 6887679 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/423,714
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-423-714-6

Query Match 99.8%; Score 3871; DB 2; Length 766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHELYLKQENNLVFNNAEYGNSSVFLNSTFDFEGHSINDYSISPD 60
DB 51 NTYRLKLYSLRWISDHELYLKQENNLVFNNAEYGNSSVFLNSTFDFEGHSINDYSISPD 110
QY 61 GQFILLYNNYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 120
DB 111 GQFILLYNNYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYLSALWSPNGTFLAYAQFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYLSALWSPNGTFLAYAQFND 230


```

Db      642 KCGIAPVSVSWYYAYVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVVDYLLINGTA 701
Qy      658 DDNVHFQOQAISKALVDVGVDFQAMWYTTDEDHGIASSTAHOIYTHMSHFIKQCFSL 715
Db      702 DDNVHFQOQAISKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLTYTHMTFLKQCFSL 758

RESULT 15
US-08-619-280A-2
; Sequence 2, Application US/08619280A
; Patent No. 5767242
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280A
; FILING DATE: 18-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5767242man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-619-280A-2

Query Match      55.1%; Score 2138; DB 1; Length 760;
Best Local Similarity 53.3%; Pred. No. 1.7e-180;
Matches 383; Conservative 134; Mismatches 187; Indels 14; Gaps 7;

Qy      2 TYRLKLSLWISDHEVLYKQ-ENNILVFNAAEYGNSSVFLNSTFDEFGHSIN--DYSIS 58
Db      51 TFSYKTFPPNWSIQEQYELHOSADNNIVLYNIETGQSYTILSNRTM----KSNVNASNYGLS 106
Qy      59 PDGQFTILLENYVKQWRHSYASDIYDLNKRQLITEERIPNNNTQWTVWSPVGHKLAYYW 118
Db      107 PDQRFVYLEDYSKLWRYSTATYIYDLNNGEVRGNELPRPIQYLWSPVGSKLAYVY 166
Qy      119 NNDIYVKIENPLPSYRITWTGKEDIYNGITDWWYEEVESAYSALWSPNGTFLAYAQF 178
Db      167 QNNIYLKQRGDPFPQITFNGRENKIFNGIPDWYBEEMLPKTKYALWWSNPKFLAYAEF 226
Qy      179 NDFEVLPIEYFYSDESLOYPKTVRPYKAGAVNTPVFPVNTDLSSTVNTATSIQIT 238
Db      227 NDKDIPVIAYSYGDE--QYPRINIPYKAGAKNPVIRIFIDTTPAYVGPQ---EVP 281
Qy      239 APASMLTGHYLQCDVTWATERISLOWLRRIQNSYVNDICDYDESSGRMNCLVARQHIEM 298
Db      282 VPAMIASSDYFSLWTLVWTDVRCVQLWLRKQVNSVLSCDPREDWQTDWDCPKTQEHIE 341

```

```

Qy      299 STTGWVGRFRPSEPHTLDGNSFYKLIISNEBGRHICYFQIDKKDCTFTIKGTWVEVIGIE 358
Db      342 SRTGMAGGFFVRPVSFYDAISYYKIFSDKDGKGIHYIKDTVVENAIQITSGKWEAINIF 401
Qy      359 ALTSYLYYISNEYKMGMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYVSFSFKEAKY 417
Db      402 RVTQDSLFSYSSNEFEYFGRRIYRISIGSYPPSKCVCTCHLKERCCQYVTASFYAKY 461
Qy      418 YQRCSGPGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQ 477
Db      462 YALVCYGPPISTLHDGRDQEIKILEENKELENALKNIQLPKBEIKKLEVDIEITLWYK 521
Qy      478 MILPHFDKSKYPLLLDYVAGPCSQKADTVFLRWATYLASTENIIVASFDRGSGSYQG 537
Db      522 MILPQFDRSKYPLLIQYVGGPCSQSVRSVFAVNMWISYLASKEGMVIALVDGRGTAFOG 581
Qy      538 DKIMHAINRRLGTFEVEDQIEAAROPSKMGFVDNKRKRIATWGSYGGYVTSMLVSGSGGVF 597
Db      582 DKLTAYVRKLGVEVEVDQITAVRKFIEMGFIDEKKRIALWGSYGGYVSVSSLASLSTGLF 641
Qy      598 KCGIAPVSVSWYYAYVYTERFMGLPTPEDNLDHYNSTVMSRAENFKQVEYLLIHGTA 657
Db      642 KCGIAPVSVSWYYAYVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVVDYLLINGTA 701
Qy      658 DDNVHFQOQAISKALVDVGVDFQAMWYTTDEDHGIASSTAHOIYTHMSHFIKQCFSL 715
Db      702 DDNVHFQOQAISKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLTYTHMTFLKQCFSL 758

```

Search completed: January 27, 2006, 23:59:03
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 23:57:48 ; Search time 67 Seconds
(without alignments)
4465.161 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766
Perfect score: 3877
Sequence: 1 NTVRLKLYSLRWISDHEYLX.....AHQHIYTHMSHFQKCFSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	3877	100.0	766	3	US-09-265-606-3
2	3877	100.0	766	3	US-09-993-959-1
3	3877	100.0	766	4	US-10-295-027-590
4	3877	100.0	766	5	US-10-952-459-18
5	3877	100.0	766	5	US-10-631-467-565
6	3871	99.8	766	4	US-10-002-593-6
7	3871	99.8	766	4	US-10-165-603-7
8	3871	99.8	766	4	US-10-423-714-6
9	3871	99.8	766	4	US-10-295-027-922
10	3871	99.8	766	4	US-10-794-899-41
11	3871	99.8	766	6	US-11-041-674-6
12	3867	99.7	766	5	US-10-476-264-147
13	3867	99.7	766	5	US-10-770-712-1
14	3358.5	86.6	767	4	US-10-770-712-2
15	3347.5	86.3	767	4	US-10-165-603-4
16	3347.5	86.3	767	4	US-10-794-899-38
17	3340	86.1	760	5	US-10-631-467-1390
18	3324	85.7	760	5	US-10-770-712-3
19	2145	55.3	760	5	US-10-723-860-4171
20	2138	55.1	760	3	US-09-265-606-2
21	2138	55.1	760	4	US-10-177-293-136
22	2138	55.1	760	4	US-10-301-822-55
23	2138	55.1	760	5	US-10-884-070A-13
24	1276.5	32.9	504	4	US-10-072-012-863
25	1221	31.5	228	5	US-10-476-264-114
26	1217	31.4	746	5	US-10-476-264-69
27	1217	31.4	746	5	US-10-476-264-70

28	1217	31.4	746	5	US-10-476-264-72	Sequence 72, Appl
29	1217	31.4	789	5	US-10-476-264-71	Sequence 71, Appl
30	1217	31.4	796	3	US-09-976-674-5	Sequence 5, Appl
31	1217	31.4	796	3	US-09-870-133-2	Sequence 2, Appl
32	1217	31.4	796	4	US-10-160-501-5	Sequence 5, Appl
33	1217	31.4	796	5	US-10-982-512-5	Sequence 5, Appl
34	1217	31.4	796	5	US-10-433-757-2	Sequence 2, Appl
35	1217	31.4	796	5	US-10-476-264-68	Sequence 68, Appl
36	1217	31.4	796	5	US-10-476-264-106	Sequence 106, App
37	1217	31.4	796	5	US-10-476-264-110	Sequence 110, App
38	1217	31.4	796	5	US-10-476-264-142	Sequence 142, App
39	1217	31.4	796	5	US-10-476-264-146	Sequence 146, App
40	1217	31.4	797	5	US-10-476-264-36	Sequence 36, Appl
41	1211	31.2	798	4	US-10-210-130-120	Sequence 120, App
42	1198	30.9	743	4	US-10-363-937-4	Sequence 4, Appl
43	1198	30.9	743	6	US-11-187-040-4	Sequence 4, Appl
44	1196	30.8	706	3	US-09-976-674-41	Sequence 41, Appl
45	1196	30.8	706	5	US-10-982-512-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-09-265-606-3
; Sequence 3, Application US/09265606
; Patent No. US20020034789A1
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,606
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/619,280
FILING DATE: 18-MARCH-1996
APPLICATION NUMBER: 08/230,491
FILING DATE: 20-APRIL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20020034789Alman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LJD 5330.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-09-265-606-3
Query Match 100.0%; Score 3877; DB 3; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NTVRLKLYSLRWISDHEYLKQENNLVFNAYGNSVFLNSTDFFGHSINDYSISPD 60

Db 51 NTRRLKLYSLRWISDHELYLKQENNLVFNABYGNSSVFLNSTFDFEGHSINDYSISPD 110
QY 61 GQFILLEYNVVKQWRHSYTTASDYIDLNKRQLITEERIIPNNTQVTTWSPVGHKLAAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTTASDYIDLNKRQLITEERIIPNNTQVTTWSPVGHKLAAYVWNN 170
QY 121 DIYVKIEPNLPYSRIITWTGKEDIINYGITDWTVEEVFSAYSALMWSPNGTFFLAYAQFND 180
Db 171 DIYVKIEPNLPYSRIITWTGKEDIINYGITDWTVEEVFSAYSALMWSPNGTFFLAYAQFND 230
QY 181 TEVPLIEYSFYSDLESLOYPKTVRPYPKAGAVNPTVKFFVNTDSLSVVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLESLOYPKTVRPYPKAGAVNPTVKFFVNTDSLSVVTNATSIQITAP 290
QY 241 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCILVARQHIEMST 300
Db 291 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCILVARQHIEMST 350
QY 301 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHI CYFQIDKKDCTFTTKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHI CYFQIDKKDCTFTTKGTWEVIGIEAL 410
QY 361 TSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
Db 411 TSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 470
QY 421 RCGSPGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQMI 480
Db 471 RCGSPGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQMI 530
QY 481 PPHFDKSKYPLLLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDGRGSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDGRGSGYQGDKI 590
QY 541 MHAINRRLCTFEVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGVVTSMVLGSGGVFKCG 600
Db 591 MHAINRRLCTFEVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGVVTSMVLGSGGVFKCG 650
QY 601 IAVAPVSRWEYDVSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 2
US-09-993-959-1
; Sequence 1, Application US/09993959
; Publication No. US20030165489A1
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/993,959
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-993-959-1
Query Match 100.0%; Score 3877; DB 3; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NTRRLKLYSLRWISDHELYLKQENNLVFNABYGNSSVFLNSTFDFEGHSINDYSISPD 60

Db 51 NTRRLKLYSLRWISDHELYLKQENNLVFNABYGNSSVFLNSTFDFEGHSINDYSISPD 110
QY 61 GQFILLEYNVVKQWRHSYTTASDYIDLNKRQLITEERIIPNNTQVTTWSPVGHKLAAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTTASDYIDLNKRQLITEERIIPNNTQVTTWSPVGHKLAAYVWNN 170
QY 121 DIYVKIEPNLPYSRIITWTGKEDIINYGITDWTVEEVFSAYSALMWSPNGTFFLAYAQFND 180
Db 171 DIYVKIEPNLPYSRIITWTGKEDIINYGITDWTVEEVFSAYSALMWSPNGTFFLAYAQFND 230
QY 181 TEVPLIEYSFYSDLESLOYPKTVRPYPKAGAVNPTVKFFVNTDSLSVVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLESLOYPKTVRPYPKAGAVNPTVKFFVNTDSLSVVTNATSIQITAP 290
QY 241 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCILVARQHIEMST 300
Db 291 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCILVARQHIEMST 350
QY 301 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHI CYFQIDKKDCTFTTKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHI CYFQIDKKDCTFTTKGTWEVIGIEAL 410
QY 361 TSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
Db 411 TSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 470
QY 421 RCGSPGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQMI 480
Db 471 RCGSPGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQMI 530
QY 481 PPHFDKSKYPLLLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDGRGSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDGRGSGYQGDKI 590
QY 541 MHAINRRLCTFEVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGVVTSMVLGSGGVFKCG 600
Db 591 MHAINRRLCTFEVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGVVTSMVLGSGGVFKCG 650
QY 601 IAVAPVSRWEYDVSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 3
US-10-295-027-590
; Sequence 590, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394

;
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 590
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-590

Query Match 100.0%; Score 3877; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	NTYRLKLYSLRWISDHEHYLKQENNILVFNAYGNSVFLNSTDFFDFGHSINDYSISPD	60
Db	51	NTYRLKLYSLRWISDHEHYLKQENNILVFNAYGNSVFLNSTDFFDFGHSINDYSISPD	110
Qy	61	GOFTLLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVWSPVGHKLAAYVNN	120
Db	111	GOFTLLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVWSPVGHKLAAYVNN	170
Qy	121	DIYVKIEPNLPSYRITWTGKEDIYNGITDMVYEEVFSAYSAWMSPNGTFLAYAQFND	180
Db	171	DIYVKIEPNLPSYRITWTGKEDIYNGITDMVYEEVFSAYSAWMSPNGTFLAYAQFND	230
Qy	181	TEVPLIEYSFSDSLOQPKTVRVPYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQTAP	240
Db	231	TEVPLIEYSFSDSLOQPKTVRVPYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQTAP	290
Qy	241	ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST	300
Db	291	ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST	350
Qy	301	TGWVGRFRPSEPFPHTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWIEVIGIEAL	360
Db	351	TGWVGRFRPSEPFPHTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWIEVIGIEAL	410
Qy	361	TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL	420
Db	411	TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL	470
Qy	421	RCSGPGGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFTILNETKFWYOMIL	480
Db	471	RCSGPGGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFTILNETKFWYOMIL	530
Qy	481	PPHFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI	540
Db	531	PPHFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI	590
Qy	541	MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTVSMVLGSSGVPKCG	600
Db	591	MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTVSMVLGSSGVPKCG	650
Qy	601	IAPAPSRWEYDVSVYTERYMGFLPTPEDNLHDYRNSVTMSRAENFKQVEYLLIHGTADDN	660
Db	651	IAPAPSRWEYDVSVYTERYMGFLPTPEDNLHDYRNSVTMSRAENFKQVEYLLIHGTADDN	710
Qy	661	VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFIIKQCFSLP	716

Db 711 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFIIKQCFSLP 766
RESULT 4
US-10-952-459-18
; Sequence 18, Application US/10952459
; Publication No. US20050074805A1
; GENERAL INFORMATION:
; APPLICANT: Kochan, Jarema Peter
; APPLICANT: Martin, Mitchell Lee
; APPLICANT: Robinski, James Andrew
; TITLE OF INVENTION: Specific Markers for Diabetes
; FILE REFERENCE: 21270US1
; CURRENT APPLICATION NUMBER: US/10/952,459
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-952-459-18

Query Match 100.0%; Score 3877; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	NTYRLKLYSLRWISDHEHYLKQENNILVFNAYGNSVFLNSTDFFDFGHSINDYSISPD	60
Db	51	NTYRLKLYSLRWISDHEHYLKQENNILVFNAYGNSVFLNSTDFFDFGHSINDYSISPD	110
Qy	61	GOFTLLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVWSPVGHKLAAYVNN	120
Db	111	GOFTLLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVWSPVGHKLAAYVNN	170
Qy	121	DIYVKIEPNLPSYRITWTGKEDIYNGITDMVYEEVFSAYSAWMSPNGTFLAYAQFND	180
Db	171	DIYVKIEPNLPSYRITWTGKEDIYNGITDMVYEEVFSAYSAWMSPNGTFLAYAQFND	230
Qy	181	TEVPLIEYSFSDSLOQPKTVRVPYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQTAP	240
Db	231	TEVPLIEYSFSDSLOQPKTVRVPYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQTAP	290
Qy	241	ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST	300
Db	291	ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST	350
Qy	301	TGWVGRFRPSEPFPHTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWIEVIGIEAL	360
Db	351	TGWVGRFRPSEPFPHTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWIEVIGIEAL	410
Qy	361	TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL	420
Db	411	TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL	470
Qy	421	RCSGPGGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFTILNETKFWYOMIL	480
Db	471	RCSGPGGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFTILNETKFWYOMIL	530
Qy	481	PPHFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI	540
Db	531	PPHFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI	590
Qy	541	MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTVSMVLGSSGVPKCG	600
Db	591	MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTVSMVLGSSGVPKCG	650
Qy	601	IAPAPSRWEYDVSVYTERYMGFLPTPEDNLHDYRNSVTMSRAENFKQVEYLLIHGTADDN	660
Db	651	IAPAPSRWEYDVSVYTERYMGFLPTPEDNLHDYRNSVTMSRAENFKQVEYLLIHGTADDN	710
Qy	661	VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFIIKQCFSLP	716

Db 711 VHFQSAQISKALVDGVDFQAMWYTDHGIASSTAHOHIYTHMSHFHKQCFSLP 766
|||||
RESULT 5
US-10-631-467-565
; Sequence 565, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for broncheal asthma, or chronic obstructive p
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 565
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-467-565

Query Match 100.0%; Score 3877; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTRYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLENSTDFDFGHSHINDYISPD 60
|||||
Db 51 NTRYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLENSTDFDFGHSHINDYISPD 110
|||||
Qy 61 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVTWS PVGHKLA YVWNN 120
|||||
Db 111 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVTWS PVGHKLA YVWNN 170
|||||
Qy 121 DIYVKIEPNLPYRITWTGKEDI IYNGITDWMYEEVFSA YSALWSPNGTFLAYAQFND 180
|||||
Db 171 DIYVKIEPNLPYRITWTGKEDI IYNGITDWMYEEVFSA YSALWSPNGTFLAYAQFND 230
|||||
Qy 181 TEVPLIEYSFYSDESLOYPKTVRVPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 240
|||||
Db 231 TEVPLIEYSFYSDESLOYPKTVRVPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 290
|||||
Qy 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
|||||
Db 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
|||||
Qy 301 TGWGRFRPSEPHTLDGNSFYKII SNEEGYRHI CYFQIDKKDCTFITKGTWEVIGIEAL 360
|||||
Db 351 TGWGRFRPSEPHTLDGNSFYKII SNEEGYRHI CYFQIDKKDCTFITKGTWEVIGIEAL 410
|||||
Qy 361 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSEAKYQYL 420
|||||
Db 411 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSEAKYQYL 470
|||||
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMP SKKLDPIILNETKFWQMIL 480
|||||
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMP SKKLDPIILNETKFWQMIL 530
|||||
Qy 481 PPHFDKSKYPLLLDDVYAGCSQKADTVFRLNATYLASTENI IIVASFDFGRSGYQGDKI 540
|||||
Db 531 PPHFDKSKYPLLLDDVYAGCSQKADTVFRLNATYLASTENI IIVASFDFGRSGYQGDKI 590
|||||
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSWVLGSGSGYVFKCG 600
|||||
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSWVLGSGSGYVFKCG 650
|||||
Qy 601 IAVAPVSRWEYDVSVYTERYMG LPTPEDNL DHRNSTVMSRAENFKQVEYLLIHGTADDN 660
|||||

Db 651 IAVAPVSRWEYDVSVYTERYMG LPTPEDNL DHRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDGVDFQAMWYTDHGIASSTAHOHIYTHMSHFHKQCFSLP 716
|||||
Db 711 VHFQSAQISKALVDGVDFQAMWYTDHGIASSTAHOHIYTHMSHFHKQCFSLP 766
|||||
RESULT 6
US-10-002-593-6
; Sequence 6, Application US/10002593
; Publication No. US20020137120A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTI
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Acty Docket No. US20020137120A1 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-593-6

Query Match 99.8%; Score 3871; DB 4; Length 766;
Best Local Similarity 99.9%; Pred. No. 1.8e-311;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTRYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLENSTDFDFGHSHINDYISPD 60
|||||
Db 51 NTRYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLENSTDFDFGHSHINDYISPD 110
|||||
Qy 61 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVTWS PVGHKLA YVWNN 120
|||||
Db 111 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVTWS PVGHKLA YVWNN 170
|||||
Qy 121 DIYVKIEPNLPYRITWTGKEDI IYNGITDWMYEEVFSA YSALWSPNGTFLAYAQFND 180
|||||
Db 171 DIYVKIEPNLPYRITWTGKEDI IYNGITDWMYEEVFSA YSALWSPNGTFLAYAQFND 230
|||||
Qy 181 TEVPLIEYSFYSDESLOYPKTVRVPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 240
|||||
Db 231 TEVPLIEYSFYSDESLOYPKTVRVPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 290
|||||
Qy 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
|||||
Db 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
|||||
Qy 301 TGWGRFRPSEPHTLDGNSFYKII SNEEGYRHI CYFQIDKKDCTFITKGTWEVIGIEAL 360
|||||
Db 351 TGWGRFRPSEPHTLDGNSFYKII SNEEGYRHI CYFQIDKKDCTFITKGTWEVIGIEAL 410
|||||
Qy 361 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSEAKYQYL 420
|||||
Db 411 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSEAKYQYL 470
|||||
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMP SKKLDPIILNETKFWQMIL 480
|||||
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMP SKKLDPIILNETKFWQMIL 530
|||||
Qy 481 PPHFDKSKYPLLLDDVYAGCSQKADTVFRLNATYLASTENI IIVASFDFGRSGYQGDKI 540
|||||
Db 531 PPHFDKSKYPLLLDDVYAGCSQKADTVFRLNATYLASTENI IIVASFDFGRSGYQGDKI 590
|||||
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSWVLGSGSGYVFKCG 600
|||||
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSWVLGSGSGYVFKCG 650
|||||

Qy 601 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHYTHMSHPFKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHYTHMSHPFKQCFSLP 766

RESULT 7

US-10-165-603-7

; Sequence 7, Application US/10165603

; Publication No. US20030021792A1

; GENERAL INFORMATION:

; APPLICANT: Roben, Paul W.

; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE

; TITLE OF INVENTION: PROTEINS

; FILE REFERENCE: TPTECH.001A

; CURRENT APPLICATION NUMBER: US/10/165,603

; CURRENT FILING DATE: 2002-06-07

; PRIOR APPLICATION NUMBER: 60/297,021

; PRIOR FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: 60/305,117

; PRIOR FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 766

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-165-603-7

Query Match 99.8%; Score 3871; DB 4; Length 766;

Best Local Similarity 99.9%; Pred. No. 1.8e-311;

Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTYRLKLYLRWISDHEHYLYKQENNLVFNAYGNSVFLNSTDFFDFGHSINDYSISPD 60
Db 51 NTYRLKLYLRWISDHEHYLYKQENNLVFNAYGNSVFLNSTDFFDFGHSINDYSISPD 110

Qy 61 GQFILLEYNVYKQWRHSYTASYDIYDLNKRQLITEERI PNTQVWTVSPVGHKLAYWNN 120
Db 111 GQFILLEYNVYKQWRHSYTASYDIYDLNKRQLITEERI PNTQVWTVSPVGHKLAYWNN 170

Qy 121 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWVYEEVFSAYSAWNSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWVYEEVFSAYSAWNSPNGTFLAYAQFND 230

Qy 181 TEVPLIEYSFYSDESLOYPKTVRPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDESLOYPKTVRPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290

Qy 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIEWST 300
Db 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIEWST 350

Qy 301 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFTTKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFTTKGTWEVIGIEAL 410

Qy 361 TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470

Qy 421 RCGSGGLPLYTLTHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 480
Db 471 RCGSGGLPLYTLTHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 530

Qy 481 PPHFDKSKYPLLLDVTYAGPCSQADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDVTYAGPCSQADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 590

Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGVYVTSWVLGSGGVKPKG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGVYVTSWVLGSGGVKPKG 650
Qy 601 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHYTHMSHPFKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHYTHMSHPFKQCFSLP 766

RESULT 8

US-10-423-714-6

; Sequence 6, Application US/10423714

; Publication No. US20030180828A1

; GENERAL INFORMATION:

; APPLICANT: Vanderbilt University

; APPLICANT: Brown, Nancy J.

; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTI

; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA

; FILE REFERENCE: Acty Docket No. US20030180828A1 1242/48/2

; CURRENT APPLICATION NUMBER: US/10/423,714

; CURRENT FILING DATE: 2003-04-25

; PRIOR APPLICATION NUMBER: 60/244,524

; PRIOR FILING DATE: 2000-10-31

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 766

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-423-714-6

Query Match 99.8%; Score 3871; DB 4; Length 766;

Best Local Similarity 99.9%; Pred. No. 1.8e-311;

Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTYRLKLYLRWISDHEHYLYKQENNLVFNAYGNSVFLNSTDFFDFGHSINDYSISPD 60
Db 51 NTYRLKLYLRWISDHEHYLYKQENNLVFNAYGNSVFLNSTDFFDFGHSINDYSISPD 110

Qy 61 GQFILLEYNVYKQWRHSYTASYDIYDLNKRQLITEERI PNTQVWTVSPVGHKLAYWNN 120
Db 111 GQFILLEYNVYKQWRHSYTASYDIYDLNKRQLITEERI PNTQVWTVSPVGHKLAYWNN 170

Qy 121 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWVYEEVFSAYSAWNSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWVYEEVFSAYSAWNSPNGTFLAYAQFND 230

Qy 181 TEVPLIEYSFYSDESLOYPKTVRPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDESLOYPKTVRPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290

Qy 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIEWST 300
Db 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIEWST 350

Qy 301 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFTTKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFTTKGTWEVIGIEAL 410

Qy 361 TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470

Qy 421 RCGSGGLPLYTLTHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 480
Db 471 RCGSGGLPLYTLTHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 530

Qy 481 PPHFDKSKYPLLLDVTYAGPCSQADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDVTYAGPCSQADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 590

Db 531 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 590
 QY 541 MHAINRLGTFEVEDQTEAARQSKMGFVNDKRIAIWGSYGGYVTSWVLGSGSGVPKCG 600
 Db 591 MHAINRLGTFEVEDQTEAARQSKMGFVNDKRIAIWGSYGGYVTSWVLGSGSGVPKCG 650
 QY 601 IAVAPVSRWEYVDSVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
 Db 651 IAVAPVSRWEYVDSVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
 QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
 Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 9
 US-10-295-027-922
 ; Sequence 922, Application US/10295027
 ; Publication No. US20030232350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; FILE REFERENCE: 018501-0125000S
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 922
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-295-027-922

Query Match 99.8%; Score 3871; DB 4; Length 766;
 Best Local Similarity 99.9%; Pred. No. 1.8e-311;
 Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NTRRLKLYSRWISDHELYLKQENNLVFNAEYGNSSVFLENSTFDFGHSINDYISPD 60
 Db 51 NTRRLKLYSRWISDHELYLKQENNLVFNAEYGNSSVFLENSTFDFGHSINDYISPD 110
 QY 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 120

Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 170
 QY 121 DIYVKIEPMLPSYRITWTCKEDIIYNGITDWDYVEEVFSAYSALWMSPNGTFLAYAQFND 180
 Db 171 DIYVKIEPMLPSYRITWTCKEDIIYNGITDWDYVEEVFSAYSALWMSPNGTFLAYAQFND 230
 QY 181 TEVPLIEYFDESLOYPKTVRVPYPKAGANPTVKFFVNTDLSLSSVTNATSIQITAP 240
 Db 231 TEVPLIEYFDESLOYPKTVRVPYPKAGANPTVKFFVNTDLSLSSVTNATSIQITAP 290
 QY 241 ASWMLIGDHVLCVDTWATQERISLQWLRRIQNTSVMDICDYDESSGRWNCVLAHQHLEMS 300
 Db 291 ASWMLIGDHVLCVDTWATQERISLQWLRRIQNTSVMDICDYDESSGRWNCVLAHQHLEMS 350
 QY 301 TGVVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFTTKGTWEVIGIAL 360
 Db 351 TGVVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFTTKGTWEVIGIAL 410
 QY 361 TSDYLYYISNEYKMGPGGRNLYKIQISDYTKVTCLSCELNPERCOYYSVSFSKEAKYQL 420
 Db 411 TSDYLYYISNEYKMGPGGRNLYKIQIDYTKVTCLSCELNPERCOYYSVSFSKEAKYQL 470
 QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFIILNETKFWTQOMIL 480
 Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFIILNETKFWTQOMIL 530
 QY 481 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 540
 Db 531 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 590
 QY 541 MHAINRLGTFEVEDQTEAARQSKMGFVNDKRIAIWGSYGGYVTSWVLGSGSGVPKCG 600
 Db 591 MHAINRLGTFEVEDQTEAARQSKMGFVNDKRIAIWGSYGGYVTSWVLGSGSGVPKCG 650
 QY 601 IAVAPVSRWEYVDSVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
 Db 651 IAVAPVSRWEYVDSVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
 QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
 Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 10
 US-10-794-899-41
 ; Sequence 41, Application US/10794899
 ; Publication No. US20040146516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Utah Ventures
 ; TITLE OF INVENTION: Lumen-Exposed Molecules and Methods Targeted Delivery
 ; FILE REFERENCE: 27110-715
 ; CURRENT APPLICATION NUMBER: US/10/794,899
 ; CURRENT FILING DATE: 2004-03-05
 ; NUMBER OF SEQ ID NOS: 121
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 41
 ; LENGTH: 766
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-794-899-41

Query Match 99.8%; Score 3871; DB 4; Length 766;
 Best Local Similarity 99.9%; Pred. No. 1.8e-311;
 Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NTRRLKLYSRWISDHELYLKQENNLVFNAEYGNSSVFLENSTFDFGHSINDYISPD 60
 Db 51 NTRRLKLYSRWISDHELYLKQENNLVFNAEYGNSSVFLENSTFDFGHSINDYISPD 110
 QY 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 120
 Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 170

QY 121 DIYVKIEPNLPSYRITWTGKEDIIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 230
QY 181 TEVPLIEYSFSDLSQYPKTVRPYPKAGAVNTPKFFVNTDLSSTVNTNATSIQITAP 240
DB 231 TEVPLIEYSFSDLSQYPKTVRPYPKAGAVNTPKFFVNTDLSSTVNTNATSIQITAP 290
QY 241 ASMLIGDHYLCDVTWATQERISLQWLRRIONYSVMDICDYDESSGRNCLVAROHIE MST 300
DB 291 ASMLIGDHYLCDVTWATQERISLQWLRRIONYSVMDICDYDESSGRNCLVAROHIE MST 350
QY 301 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVIGIEAL 360
DB 351 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVIGIEAL 410
QY 361 TSDVLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQ 420
DB 411 TSDVLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQ 470
QY 421 RCSGPGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFIILNETKFWYQ 480
DB 471 RCSGPGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFIILNETKFWYQ 530
QY 481 PPFDKSKYPLLLDVVAGPCSKADTVFRLNWTYLASTENIIIVASFDGRGSGYQGDKI 540
DB 531 PPFDKSKYPLLLDVVAGPCSKADTVFRLNWTYLASTENIIIVASFDGRGSGYQGDKI 590
QY 541 MHAINRLRGTEVEDEQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 600
DB 591 MHAINRLRGTEVEDEQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 650
QY 601 IAVAPSRWEYSDSVYTERYNGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPSRWEYSDSVYTERYNGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 11
US-11-041-674-6
; Sequence 6, Application US/11041674
; Publication No. US20050181468A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN
; FILE REFERENCE: Acty Docket No. 1242/48/2/2/2
; CURRENT APPLICATION NUMBER: US/11/041,674
; CURRENT FILING DATE: 2005-01-24
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 10/002,593
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 10/423,714
; PRIOR FILING DATE: 2003-4-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-041-674-6
Query Match 99.8%; Score 3871; DB 6; Length 766;
Best Local Similarity 99.9%; Pred. No. 1.8e-311;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 NTVRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSFTDFGHSINDYSISPD 60

DB 51 NTVRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSFTDFGHSINDYSISPD 110
QY 61 GQFTLEEVYVVKWRHSYTASYDIYDLNKEQLITEERI PNTTQWVWTSFVGHKLAYVWNN 120
DB 111 GQFTLEEVYVVKWRHSYTASYDIYDLNKEQLITEERI PNTTQWVWTSFVGHKLAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 230
QY 181 TEVPLIEYSFSDLSQYPKTVRPYPKAGAVNTPKFFVNTDLSSTVNTNATSIQITAP 240
DB 231 TEVPLIEYSFSDLSQYPKTVRPYPKAGAVNTPKFFVNTDLSSTVNTNATSIQITAP 290
QY 241 ASMLIGDHYLCDVTWATQERISLQWLRRIONYSVMDICDYDESSGRNCLVAROHIE MST 300
DB 291 ASMLIGDHYLCDVTWATQERISLQWLRRIONYSVMDICDYDESSGRNCLVAROHIE MST 350
QY 301 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVIGIEAL 360
DB 351 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVIGIEAL 410
QY 361 TSDVLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQ 420
DB 411 TSDVLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQ 470
QY 421 RCSGPGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFIILNETKFWYQ 480
DB 471 RCSGPGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFIILNETKFWYQ 530
QY 481 PPFDKSKYPLLLDVVAGPCSKADTVFRLNWTYLASTENIIIVASFDGRGSGYQGDKI 540
DB 531 PPFDKSKYPLLLDVVAGPCSKADTVFRLNWTYLASTENIIIVASFDGRGSGYQGDKI 590
QY 541 MHAINRLRGTEVEDEQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 600
DB 591 MHAINRLRGTEVEDEQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 650
QY 601 IAVAPSRWEYSDSVYTERYNGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPSRWEYSDSVYTERYNGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 12
US-10-476-264-147
; Sequence 147, Application US/10476264
; Publication No. US20050123910A1
; GENERAL INFORMATION:
; APPLICANT: Cookson, William Osmond Charles Michael
; APPLICANT: Moffat, Miriam Fleur
; APPLICANT: Allen, Maxine
; APPLICANT: Lench, Nick
; TITLE OF INVENTION: Enzyme and SNP marker for disease
; FILE REFERENCE: 16721-002US1
; CURRENT APPLICATION NUMBER: US/10/476,264
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PCT/GB02/01887
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: GB0110044.5
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0110046.0
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0124594.3
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: GB0124575.2
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 421
; SOFTWARE: PatentIn version 3.1


```
; SEQ ID NO 147
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-476-264-147

Query Match          99.7%; Score 3867; DB 5; Length 766;
Best Local Similarity 99.7%; Pred. No. 3.9e-311;
Matches 714; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTPDEFGHSINDYISPD 60
Db 51 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTPDEFGHSINDYISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWVYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWVYEEVFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 290
Qy 241 ASMLIGDHVLCVDTWATQERISLOWLRRIQNTSVMDICDYDESSGRWNCCLVARQHIE MST 300
Db 291 ASMLIGDHVLCVDTWATQERISLOWLRRIQNTSVMDICDYDESSGRWNCCLVARQHIE MST 350
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFTFTKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFTFTKGTWEVIGIEAL 410
Qy 361 TSDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWQMIL 530
Qy 481 PPHPDSSKKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDGRSGGYQGDKI 540
Db 531 PPHPDSSKKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDGRSGGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGWSYGGYVTSMLVLSGSGGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGWSYGGYVTSMLVLSGSGGVFKCG 650
Qy 601 IAVAPVSRWEYYSVYTERVNGLPTEPDNLDHVRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYYSVYTERVNGLPTEPDNLDHVRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFHKQCFSLP 716
Db 711 VHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFHKQCFSLP 766
```

```
RESULT 13
US-10-770-712-1
; Sequence 1, Application US/10770712
; Publication No. US20050170333A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
; FILE REFERENCE: IMSC12.008A
; CURRENT APPLICATION NUMBER: US/10/770,712
; CURRENT FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 766
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-712-1

Query Match          99.7%; Score 3867; DB 5; Length 766;
Best Local Similarity 99.7%; Pred. No. 3.9e-311;
Matches 714; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTPDEFGHSINDYISPD 60
Db 51 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTPDEFGHSINDYISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWVYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWVYEEVFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 290
Qy 241 ASMLIGDHVLCVDTWATQERISLOWLRRIQNTSVMDICDYDESSGRWNCCLVARQHIE MST 300
Db 291 ASMLIGDHVLCVDTWATQERISLOWLRRIQNTSVMDICDYDESSGRWNCCLVARQHIE MST 350
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFTFTKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFTFTKGTWEVIGIEAL 410
Qy 361 TSDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWQMIL 530
Qy 481 PPHPDSSKKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDGRSGGYQGDKI 540
Db 531 PPHPDSSKKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDGRSGGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGWSYGGYVTSMLVLSGSGGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGWSYGGYVTSMLVLSGSGGVFKCG 650
Qy 601 IAVAPVSRWEYYSVYTERVNGLPTEPDNLDHVRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYYSVYTERVNGLPTEPDNLDHVRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFHKQCFSLP 716
Db 711 VHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFHKQCFSLP 766
```

```
RESULT 14
US-10-770-712-2
; Sequence 2, Application US/10770712
; Publication No. US20050170333A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
; FILE REFERENCE: IMSC12.008A
; CURRENT APPLICATION NUMBER: US/10/770,712
; CURRENT FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Rattus norvegicus
```


US-10-770-712-2

Query Match 86.6%; Score 3358.5; DB 5; Length 767;
Best Local Similarity 84.8%; Pred. No. 5.4e-269;
Matches 609; Conservative 51; Mismatches 55; Indels 3;

Qy	1	NTYRLKLSLRLWISDSEHYLYIQENNILYFNAEYGNSSVFLNSTDFDEFGHSINDYSISPD	60
Db	49	NTFRKYSLSLWSDSEYLYIQENNILLFNAEHGNSIFLENSTFEI FGDSTISDYSVSPD	108
Qy	61	GQFTLLEVYVQWRHSYTSASVDIYDLNKRQLITEERI PNNTOWTWSPVGHKLAVVNN	120
Db	109	RLFVLELVYVQWRHSYTSASISYDLNKRQLITEEKIPNNTOWTWSEOGHKLAVVKN	168
Qy	121	DIYVKIEPILNPSYRITITGKEDIITNGITDWYEEBEVFSAYSALWWSNGTFLAYAQFND	180
Db	169	DIYVKIEPILPHLSRITSTGKENVIFNGINDWYEEBEEIFGAYSALWWSNGTFLAYAQFND	228
Qy	181	TEVPLIEYSFYSDSLSQPKTVVPYPKAGAVNPTVKSPFVNTDSLSVTNATSIQITAP	240
Db	229	TGVEPLIEYSFYSDSLSQPKTWIPIYPKAGAVNPTVKSPFVNTDSLSVTITPQITAP	288
Qy	241	ASMLIGDHYLCDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCILVARQHIEHST	300
Db	289	ASVTTGHDYLCDVAWVEDSRISLOWLRRIQNYSVMAICDYDKTILVWNCPTTREHIEHSTA	348
Qy	301	TGWGRRFRPSPHFTTLDGNSFYKILISNEEGYRHYICYFOIDKKD--CTFITKGTWEVIGI	357
Db	349	TGWGRRFRPAEPHFTSDGSSFYKLVSDKGXGKHICQFOKDRKPEQVCFTITKGAVEVILSI	408
Qy	358	EALTSIDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCBLNPERCOYVSVFSKEAKY	417
Db	409	EALTSIDLYYISNEYKEMPGGRNLYKIQLTDTHTNKKCLSCDLNPERCOYVSVLSKEAKY	468
Qy	418	YQLKCSGPGPLPYLTHSSVNDKGRLVRLEDNSALDKMLQNOMPSKKLDFIILNETKFWYQ	477
Db	469	YQLGCRGPGPLPYLTHRSTDKELRVLEDNSALDKMLQDVOMPSKKLDFIILNETRFFWYQ	528
Qy	478	MILPPHPDKSKYPLLLDVIYAGPCSOQADTVFRLNWTATYLASTENIIVASDFGRSGYQG	537
Db	529	MILPPHPDKSKYPLLLDVIYAGPCSOQADAAFLRNWTATYLASTENIIVASDFGRSGYQG	588
Qy	538	DKIMHAINRRLGTPEVEDQIEAARQFSGMWGVNDKNRIAIWGWSGYGYVTSWVLGSGSGVF	597
Db	589	DKIMHAINKRLGTLEVEDQIEAARQFLKMGFVDSKRVAINWGSYGYGYVTSWVLGSGSGVF	648
Qy	598	KCGIAPAVSRWEYDYSYTERYMGLEPTPEDNLDRHNRNSTVMSRAENFKQVEYILLIHGTA	657
Db	649	KCGIAPAVSRWEYDYSYTERYMGLEPTPEDNLDRHNRNSTVMSRAENFKQVEYILLIHGTA	708
Qy	658	DDNVHFOQSAQISKALVDVGVDFQAMWYTTDBDHGIASSTAHOHYTHMSHFIKOCFSL	715
Db	709	DDNVHFOQSAQISKALVDAGVDFQAMWYTTDBDHGIASSTAHOHYTHMSHFILQOCFSL	766

RESULT 15

```

US-10-165-603-4
; Sequence 4, Application US/10165603
; Publication NO. US20030021792A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Paul W.
; APPLICANT: Stevens, Anthony C.
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: TPECH.001A
; CURRENT APPLICATION NUMBER: US/10/165,603
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/297,021
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/305,117
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ IDS: 33
; SOFTWARE: FASTSEQ for Windows Version 4.0

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2006, 14:07:15 ; Search time 135 Seconds
(without alignments)

2330.335 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877

Sequence: 1 NTVRLKLSRWISWDHEYL.....AQHIYTHMGHFKQCFSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	759	2 AAR54612	Aar54612 Delta3-9
2	3877	100.0	766	2 AAR40909	Aar40909 Sequence
3	3877	100.0	766	5 ABB08991	Abb08991 Human dip
4	3877	100.0	766	5 AAG78417	Aag78417 Human dip
5	3877	100.0	766	7 ADD27855	Add27855 Human dip
6	3877	100.0	766	7 ADD46934	Add46934 Human Pro
7	3877	100.0	766	7 ADN39272	Adn39272 Cancer/an
8	3877	100.0	766	8 ADJ83981	Adj83981 Human ful
9	3877	100.0	766	8 ADJ75313	Adj75313 Marker ge
10	3877	100.0	766	8 ADO19398	Ado19398 Human PRO
11	3877	100.0	766	8 ADO19806	Ado19806 Human PRO
12	3877	100.0	766	8 ADO171612	Ado171612 Amino aci
13	3877	100.0	766	8 ADO171644	Ado171644 Amino aci
14	3877	100.0	766	8 ABM80355	Abm80355 Tumour-as
15	3877	100.0	766	8 ADP54458	Adp54458 Human PRO
16	3877	100.0	766	8 ADU06688	Adu06688 Novel bro
17	3877	100.0	766	8 ADV25525	Adv25525 Human dip
18	3877	100.0	766	9 ADY15161	Ady15161 PRO polyp
19	3877	100.0	766	9 ADY16580	Ady16580 PRO polyp
20	3877	100.0	766	9 ADZ14038	Adz14038 Human dip
21	3877	100.0	766	9 AEB94223	Aeb94223 CD26/dipe
22	3871	99.8	736	8 ADO40240	Ado40240 Human DPP
23	3871	99.8	766	5 ABG61910	Abg61910 Prostate
24	3871	99.8	766	5 AAO15555	Aao15555 Human dip

25	3871	99.8	766	6 ABP56700	Abp56700 Human liv
26	3871	99.8	766	7 ADD14045	Add14045 Human src
27	3871	99.8	766	7 ADN39604	Adn39604 Cancer/an
28	3871	99.8	766	8 ADO19400	Ado19400 Human PRO
29	3867	99.7	766	6 ABP55629	Abp55629 Human dpp
30	3867	99.7	766	8 ADQ80365	Adq80365 Dipeptidy
31	3867	99.7	766	9 AEB77579	Aeb77579 Human dip
32	3866	99.7	766	2 AAR54611	Aar54611 Native CD
33	3779	97.5	739	2 AAR54613	Aar54613 Delta24-3
34	3441	88.8	688	8 ADO71642	Ado71642 Amino aci
35	3358.5	86.6	767	9 AEB77580	Aeb77580 Rat dipep
36	3355.5	86.5	767	3 AAB11748	Aab11748 Rat dipep
37	3354.5	86.5	767	7 ADD46932	Add46932 Rat Prote
38	3347.5	86.3	767	6 ABP56699	Abp56699 Rat liver
39	3340	86.1	760	8 ADJ76138	Adj76138 Marker ge
40	3340	86.1	760	8 ADO71646	Ado71646 Amino aci
41	3340	85.7	760	9 AEB94226	Aeb94226 Mouse CD2
42	3324	85.7	760	9 AEB77581	Aeb77581 Mouse dip
43	2948	76.0	593	2 AAR40916	Aar40916 Sequence
44	2948	76.0	593	2 AAR54614	Aar54614 Delta594-
45	2145	55.3	723	9 AEB94227	Aeb94227 Human sol

ALIGNMENTS

RESULT 1

AAR54612

ID AAR54612 standard; protein; 759 AA.

XX AAR54612;

AC

XX

25-MAR-2003 (revised)

DT 09-DEC-1994 (first entry)

XX

DE Delta3-9 CD26.

XX

Human; T cell activation antigen; CD26; analogues; deletion; soluble;

KW signal peptidase; immune-stimulating; response-stimulating; AIDS;

KW immunosuppression; AIDS-related complex.

XX Homo sapiens.

XX

Key Location/Qualifiers

FT Misc-difference 2..3

FT /note= "Position of delta3-9 deletion"

XX

PN WO9409132-A1.

XX

PD 28-APR-1994.

XX

PF 19-AUG-1993; 93WO-US007923.

XX

PR 21-AUG-1992; 92US-00934162.

XX

PA (DAND) DANA FARBER CANCER INST INC.

XX

PI Morimoto C, Schlossman S, Tanaka T;

XX WPI; 1994-151317/18.

DR

XX

PT Polypeptide fragments and analogues of CD26 and encoding nucleic acid -

XX

PT useful for stimulating immune response, e.g. for treatment of AIDS to

XX

PT counteract immunosuppressive drug, and as vaccine adjuvant.

XX

PS Claim 3; Page 49-52; 85pp; English.

XX

CC The sequences given in AAR54612-14 represents analogues of the human T

XX

CC cell activation antigen CD26 which have internal deletions. The analogues

XX

CC pref. lack residues 3-9 or 24-34. These analogues are soluble under

XX

CC physiological conditions and lack enough amino acid residues to render

XX

CC them susceptible to cleavage by signal peptidase. The peptide fragments

XX

CC and analogues are useful as immune or response- stimulating therapeutics,

CC eg. they may be used for treatment of disease conditions characterised by
CC immunosuppression, eg. AIDS or AIDS-related complex, other virally or
CC environmentally-induced conditions, and certain congenital immune
CC deficiencies. The peptides can be employed to increase immune function
CC which has been impaired by use of immunosuppressive drugs, such as certain
CC chemotherapeutic drugs. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 759 AA;

```
Query Match      100.0%; Score 3877; DB 2; Length 759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLSLRWISDHEYLKQENNILVFNAEYGNSSVFLENSTDFEFGHSINDYSISPD 60
Db 44 NTYRLKLSLRWISDHEYLKQENNILVFNAEYGNSSVFLENSTDFEFGHSINDYSISPD 103
QY 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 120
Db 104 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 163
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 164 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 223
QY 181 TEVPLIEYSYDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 240
Db 224 TEVPLIEYSYDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 283
QY 241 ASMLIGHYLCVDTWATQERISIQWLRRIONYSVMDICDYDESSGRWNCCLVARQHEMST 300
Db 284 ASMLIGHYLCVDTWATQERISIQWLRRIONYSVMDICDYDESSGRWNCCLVARQHEMST 343
QY 301 TGWGRFRPSEPHEFTLDGNSFYKIIISNEGYRHCYFQIDKDCCTFITKGTWEVIGIEAL 360
Db 344 TGWGRFRPSEPHEFTLDGNSFYKIIISNEGYRHCYFQIDKDCCTFITKGTWEVIGIEAL 403
QY 361 TSDYLYIISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCQYYSVFSKEAKYQL 420
Db 404 TSDYLYIISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCQYYSVFSKEAKYQL 463
QY 421 RCGSGPLPLYTHSSVNDKGLRVLEDSALDKLQNVQMPSKLDFTILNETKFWQMIL 480
Db 464 RCGSGPLPLYTHSSVNDKGLRVLEDSALDKLQNVQMPSKLDFTILNETKFWQMIL 523
QY 481 PPFDKSKKYPYLLLDVYAGCSQKADTVFRLNWTYLASTENIIVASFDGRGSGYQGDKI 540
Db 524 PPFDKSKKYPYLLLDVYAGCSQKADTVFRLNWTYLASTENIIVASFDGRGSGYQGDKI 583
QY 541 MHAINRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 600
Db 584 MHAINRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 643
QY 601 IAVAPSVRWYSDVSVYTERVWGLPTPEDNLHDYRNSTVMSRAENKQVEVLLIHGTADDN 660
Db 644 IAVAPSVRWYSDVSVYTERVWGLPTPEDNLHDYRNSTVMSRAENKQVEVLLIHGTADDN 703
QY 661 VHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
Db 704 VHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 759
```

RESULT 2
AAR40909
ID AAR40909 standard; protein; 766 AA.
XX
AC AAR40909;
XX
XX 25-MAR-2003 (revised)
DT 05-FEB-1994 (first entry)
XX
DE Sequence encoded by human CD26 cDNA.
XX

KW Human T cell activation antigen; monoclonal antibody Tal.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT 7..28 /label= hydrophobic
FT 29..323 /label= N-terminal glycosylated region of extracellular domain
FT /note= "8 sites for N-linked glycans"
FT 324..551 /label= Cysteine rich region of extracellular domain
FT /note= "1 N-linked glycosylation site"
FT 552..766 /label= C-terminal region of extracellular domain
FT /note= "1 N-linked glycosylation site & 1 catalytic site"
FT 627..631 /label= active site of serine protease/esterase
FT /note= "fits the consensus sequence GXSGX"
XX

PN WO9316102-A1.
XX
XX
PD 19-AUG-1993.
XX
PF 09-APR-1992; 92WO-US002892.
XX
XX
PR 06-FEB-1992; 92US-00832211.
XX
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
XX
PI Morimoto C, Schlossman SF, Tanaka T;
XX
XX
DR WPI; 1993-272827/34.
XX
XX
PT Polypeptide fragments of CD26 - are capable of disrupting binding of CD45
PT and CD26 and thus interfering with T-cell activation.
XX
XX
PS Disclosure; Page 39-43; 73pp; English.
XX
XX
CC C26 is a human T cell activation antigen originally identified by its
CC reactivity with the MAb Tal. C26 cDNA library was constructed from human
CC PHA-activated T cells using the CDM7vector. The hydrophobic N-terminal of
CC the predicted CD26 polypeptide has the characteristics of a signal
CC sequence of the type II membrane protein, which is reinforced by the
CC observation that potential N-glycosylation sites are located in the
CC carboxy side of the hydrophobic core. Therefore the N-terminal 6 AAs are
CC predicted to be cytoplasmic, the next 22 AAs are predicted to transverse
CC the cytoplasmic membrane, and the 738 C-terminal AAs constitute the
CC predicted extracellular domain. (Updated on 25-MAR-2003 to correct PN
XX field.)
SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NTYRLKLSLRWISDHEYLKQENNILVFNAEYGNSSVFLENSTDFEFGHSINDYSISPD 60
Db 51 NTYRLKLSLRWISDHEYLKQENNILVFNAEYGNSSVFLENSTDFEFGHSINDYSISPD 110
QY 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 230
QY 181 TEVPLIEYSYDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 240

Db 231 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNPTKFFVNTDLSLSSVTNATSIQTAP 290
Qy 241 ASMLIGDHYLCDVTWATERISLOWLRRIQNSVMDICDYDESSGRWNCVLARQHIEMST 300
Db 291 ASMLIGDHYLCDVTWATERISLOWLRRIQNSVMDICDYDESSGRWNCVLARQHIEMST 350
Qy 301 TGVGRFRPSPHPTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEAL 360
Db 351 TGVGRFRPSPHPTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEAL 410
Qy 361 TSDLYIYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 411 TSDLYIYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFIILNETKFWYQML 480
Db 471 RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFIILNETKFWYQML 530
Qy 481 PPFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 540
Db 531 PPFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVPKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVPKCG 650
Qy 601 IAVAPVSRWEYDYSVTYRYMGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDYSVTYRYMGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 3

ABB08991
ID ABB08991 standard; protein; 766 AA.
XX
AC ABB08991;
XX
DT 19-JUN-2002 (first entry)
XX
DE Human dipeptidyl peptidase IV.
XX
KW Human; dipeptidyl peptidase IV; antiasthmatic; antiallergic;
KW antiinflammatory.
XX
OS Homo sapiens.
XX
PN US6337069-B1.
XX
PD 08-JAN-2002.
XX
PF 28-FEB-2001; 2001US-00794236.
XX
PR 28-FEB-2001; 2001US-00794236.
XX
PA (BMRA-) BMRA CORP BV.
XX
PI Grouzmann E, Lacroix J, Monod M;
XX WPI; 2002-163235/21.
XX

XX Treating a patient for mucosal inflammation associated with rhinitis,
PT sinusitis or both, by intranasally administering a peptidase that cleaves
PT at Xaa-Pro sequences, to the patient.
XX
XX Disclosure; Col 9-14; 13pp; English.
XX
XX Thus invention relates to the treating of a patient for mucosal
CC inflammation associated with rhinitis or sinusitis, comprising
CC intranasally administering a peptidase. The peptidase is considered

CC antiasthmatic, antiallergic and antiinflammatory in its action. The
CC peptidase cleaves at Xaa-Pro sequences and is useful for treating a
CC patient for mucosal inflammation associated with rhinitis or sinusitis,
CC which is the result of allergies or asthma. This sequence represents
CC human dipeptidyl peptidase IV
XX
SQ Sequence 766 AA;
Query Match 100.0%; Score 3877; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NTVRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLENSTFDFGHSINDYSISPD 60
Db 51 NTVRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLENSTFDFGHSINDYSISPD 110
Qy 61 GQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTVSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTVSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDMVVEEVPFSAYSAWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDMVVEEVPFSAYSAWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNPTKFFVNTDLSLSSVTNATSIQTAP 240
Db 231 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNPTKFFVNTDLSLSSVTNATSIQTAP 290
Qy 241 ASMLIGDHYLCDVTWATERISLOWLRRIQNSVMDICDYDESSGRWNCVLARQHIEMST 300
Db 291 ASMLIGDHYLCDVTWATERISLOWLRRIQNSVMDICDYDESSGRWNCVLARQHIEMST 350
Qy 301 TGVGRFRPSPHPTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEAL 360
Db 351 TGVGRFRPSPHPTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEAL 410
Qy 361 TSDLYIYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 411 TSDLYIYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFIILNETKFWYQML 480
Db 471 RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFIILNETKFWYQML 530
Qy 481 PPFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 540
Db 531 PPFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVPKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVPKCG 650
Qy 601 IAVAPVSRWEYDYSVTYRYMGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDYSVTYRYMGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 4

AAG78417
ID AAG78417 standard; protein; 766 AA.
XX
AC AAG78417;
XX
DT 12-APR-2002 (first entry)
XX
DE Human dipeptidyl peptidase IV amino acid sequence.
XX
KW 21953 prolyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;

KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW anti-diabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease;
KW dipeptidyl peptidase.
OS Homo sapiens.
XX
XX WO200179473-A2.
PN
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US040483.
XX
XX 18-APR-2000; 2000US-0197508P.
PR
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
XX Meyers RA, Williamson M;
PI
XX
XX WPI; 2002-034353/04.
DR
XX
XX New polypeptides 21953, member of human prollyl oligopeptidase family,
PT useful as diagnostic targets and therapeutic agents for controlling
PT cancer, lymphoma and leukemia.
XX
XX Disclosure; Fig 3; 121pp; English.
XX
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.
CC Which is cytostatic, anti-diabetic, antiarthritic, neuroprotective,
CC antithyroid, dermatological, antipsoriatic, antiasthmatic,
CC ophthalmological, antiinflammatory, nootropic, antiparkinsonian,
CC anticonvulsant, gynaecological, vasotropic, antianginal, cardiac,
CC antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC gene therapy, expression or activity of 21953 protein modulator, it is
CC useful for identifying a compound which binds to it and can be used in
CC preventing, treating or detecting a cellular proliferative or
CC differentiative disorder. The 21953 molecules can act as novel diagnostic
CC targets and therapeutic agents for controlling disorders associated with
CC the aberrant activity or degradation of peptide hormones e.g., disorders
CC associated with cell differentiation and proliferation such as cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, demyelinating diseases, vascular disorders and
CC metabolism or pain disorders. This sequence represents the amino acid
CC sequence of human dipeptidyl peptidase IV
XX
XX Sequence 766 AA;
SQ
Query Match 100.0%; Score 3877; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NTVRLKLYLSRWISDHELYLKQBNILVFNAEYGNSSVFLNFTDFEFGHSINDYSISPD 60
Db
51 NTVRLKLYLSRWISDHELYLKQBNILVFNAEYGNSSVFLNFTDFEFGHSINDYSISPD 110
QY 61 GQFLLLEYNVVKQWRHSYTASYDIYDLNKEQLITEERI PNNTQVMTWS PVGHKLAYVWNN 120
Db
111 GQFLLLEYNVVKQWRHSYTASYDIYDLNKEQLITEERI PNNTQVMTWS PVGHKLAYVWNN 170
QY 121 DIYVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYSLWNSPNTGFLAYAQFND 180
Db
171 DIYVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYSLWNSPNTGFLAYAQFND 230
QY 181 TEVPLIEYSFYSDESLOQPKTVRVPYKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 240
Db
231 TEVPLIEYSFYSDESLOQPKTVRVPYKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 290
QY 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCVLVARQHIE MST 300

Db 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCVLVARQHIE MST 350
QY 301 TGVVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCICYFQIDKKDCTFITKGTWEVIGIEAL 360
Db 351 TGVVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCICYFQIDKKDCTFITKGTWEVIGIEAL 410
QY 361 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYSVSFSKEAKYYQL 420
Db 411 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYSVSFSKEAKYYQL 470
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDPIIINETKFTQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDPIIINETKFTQMIL 530
QY 481 PPHFDKSKKYPILLDYYAGPCSQKADTVFRLNWTYLASTENIIVASFDRGSGYQGDKI 540
Db 531 PPHFDKSKKYPILLDYYAGPCSQKADTVFRLNWTYLASTENIIVASFDRGSGYQGDKI 590
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLGSGSGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLGSGSGVFKCG 650
QY 601 IAVAPVSRWEYVDSVYTERYMGIPTEPDNLDHVRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYVDSVYTERYMGIPTEPDNLDHVRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDHGHASSTAHOIYTHMSHFHKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDHGHASSTAHOIYTHMSHFHKQCFSLP 766
RESULT 5
ADD27855
ID ADD27855 standard; protein; 766 AA.
XX
AC ADD27855;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human dipeptidyl peptidase IV (DPPIV) .
XX
KW Mucosal inflammation; rhinitis; sinusitis; exopeptidase; substance P; SP;
KW neurokinin 1 receptor; NK1 receptor; allergy; asthma; antiasthmatic; human; dipeptidyl peptidase IV; DPPIV;
KW antiinflammatory; antiasthmatic; human; dipeptidyl peptidase IV; DPPIV;
XX enzyme.
OS Homo sapiens.
XX
PN US2003165489-A1.
XX
XX 04-SEP-2003.
XX
XX 27-NOV-2001; 2001US-00993959.
XX
XX 28-FEB-2001; 2001US-00794236.
XX
PA (BMRA-) BMRA CORP BV.
XX
XX Grouzmann E, Lacroix J, Monod M;
XX
XX WPI; 2003-811386/76.
XX
PT Treatment of patient for mucosal inflammation associated with rhinitis
PT and/or sinusitis involves intranasally administering peptidase that
PT cleaves at Xaa-Pro sequences or agent inhibiting binding of Sp to
PT neurokinin 1 receptor.
XX
XX Disclosure; SEQ ID NO 1; 14pp; English.
XX
XX The present invention relates to a method of treating a patient for
CC mucosal inflammation associated with rhinitis and/or sinusitis. The
CC method comprises intranasally administering to the patient a peptidase

CC that cleaves at Xaa-Pro sequences or an agent that inhibits the binding
CC of substance P (SP) to the neurokinin 1 (NK1) receptor. The peptidase is
CC an exopeptidase, preferably selected from human dipeptidyl peptidase IV
CC (DPP-IV), human quiescent cell proline dipeptidase, human dipeptidyl
CC peptidase 8, or human attractin. The method is useful for treating a
CC patient for mucosal inflammation associated with rhinitis and/or
CC sinusitis which are the result of allergies or asthma. The invention
CC provides an effective treatment of the inflammation associated with both
CC rhinitis and sinusitis. The present sequence represents human DPP-IV.
XX
XX
SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 7; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLSLRWISDHELYLKQENNLVFNAYGNSVFNLENSTDFEFGHSINDYSISPD 60
Db 51 NTYRLKLSLRWISDHELYLKQENNLVFNAYGNSVFNLENSTDFEFGHSINDYSISPD 110
QY 61 GQFILLENNYVVKQWRHSYASDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 120
Db 111 GQFILLENNYVVKQWRHSYASDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAFND 230
QY 181 TEVPLIEYSFSDSLQPKTVRPYPKAGAVNTPKFFVNTVTSLSVTVNATSIQTAP 240
Db 231 TEVPLIEYSFSDSLQPKTVRPYPKAGAVNTPKFFVNTVTSLSVTVNATSIQTAP 290
QY 241 ASMLIGDHYLQDVWATQERISLOWLRRIQNSVMDICDYDESSGRWNCVLVAROHENST 300
Db 291 ASMLIGDHYLQDVWATQERISLOWLRRIQNSVMDICDYDESSGRWNCVLVAROHENST 350
QY 301 TGWVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYQIDKDKCTFITKGTWVIGIEAL 360
Db 351 TGWVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYQIDKDKCTFITKGTWVIGIEAL 410
QY 361 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
Db 411 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 470
QY 421 RCGSGPLPLYLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFFILNETKFWYQML 480
Db 471 RCGSGPLPLYLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFFILNETKFWYQML 530
QY 481 PPFDKSKKYPDLLDVVAGPCQKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 540
Db 531 PPFDKSKKYPDLLDVVAGPCQKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 590
QY 541 MHAENRRLGTPEVEDQIEAARQFSKMGFVDNKR IAINGWSYGGYVTVSMVLGSGGVFKCG 600
Db 591 MHAENRRLGTPEVEDQIEAARQFSKMGFVDNKR IAINGWSYGGYVTVSMVLGSGGVFKCG 650
QY 601 IAVAPSRWEYSDSYVTERYNGLTPEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPSRWEYSDSYVTERYNGLTPEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHYTHMSHFIKQCFSLP 716
Db 711 VHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHYTHMSHFIKQCFSLP 766

RESULT 6
ID ADD46934
XX ADD46934 standard; protein; 766 AA.
AC ADD46934;
XX
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)

XX Human Protein AAA52308, SEQ ID NO 12620.
DE Human; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
KW Homo sapiens.
OS Unidentified.
OS WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GPHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
PA Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
XX GENBANK; AAA52308.
DR New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
PT Example 1; Page; 1017pp; English.
PS The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 7; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLSLRWISDHELYLKQENNLVFNAYGNSVFNLENSTDFEFGHSINDYSISPD 60
Db 51 NTYRLKLSLRWISDHELYLKQENNLVFNAYGNSVFNLENSTDFEFGHSINDYSISPD 110
QY 61 GQFILLENNYVVKQWRHSYASDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 120

Db 111 GQFILLVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSVRIITWTGKEDIYINGITDMWYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSVRIITWTGKEDIYINGITDMWYEEVFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFVNTDSLSSTVNTATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFVNTDSLSSTVNTATSIQITAP 290
Qy 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCILVARQHIEMST 300
Db 291 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCILVARQHIEMST 350
Qy 301 TGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHICYFQIDKDKCTFITKGTWVIGIEAL 360
Db 351 TGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHICYFQIDKDKCTFITKGTWVIGIEAL 410
Qy 361 TSDVLYYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELPERCQYVSFSKEAKYQL 420
Db 411 TSDVLYYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELPERCQYVSFSKEAKYQL 470
Qy 421 RCGSPGLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFFILNETKFWYQML 480
Db 471 RCGSPGLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFFILNETKFWYQML 530
Qy 481 PPHFDKSKYPLLLDVYAGCSQKADTVFRLNATYLASTENIIVASFDRGSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDVYAGCSQKADTVFRLNATYLASTENIIVASFDRGSGYQGDKI 590
Qy 541 MHAINRLGTFEVEDQIEAARQFSKMGFVGNKRIATWGSYGGVVTSMVLGSGGVFKCG 600
Db 591 MHAINRLGTFEVEDQIEAARQFSKMGFVGNKRIATWGSYGGVVTSMVLGSGGVFKCG 650
Qy 601 IAVAPSRWEYDYSVYTERYMGFLPTPEDNLDHYRNSTVMGSAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPSRWEYDYSVYTERYMGFLPTPEDNLDHYRNSTVMGSAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDGVDFQAMWYTDHEDHGIASSTAHQHYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDGVDFQAMWYTDHEDHGIASSTAHQHYTHMSHFIKQCFSLP 766

RESULT 7
ADN39272 ID ADN39272 standard; protein; 766 AA.
XX AC ADN39272;
XX DT 17-JUN-2004 (first entry)
XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:590.
XX KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.
XX OS Homo sapiens.
XX WO2003042661-A2.
XX PN 22-MAY-2003.
XX PD 13-NOV-2002; 2002WO-US036810.
XX PF 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.

PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI: 2003-468649/44.
DR N-PSDB; ADN39271.
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX Claim 12; SEQ ID NO 590; 1385pp; English.
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides, and
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX SQ Sequence 766 AA;
Query Match 100.0%; Score 3877; DB 7; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NTYRLKLYSLRWISDHELYLKQENNLVFNAYGNSVFLNSTDFEGCHSINDYSISPD 60
Db 51 NTYRLKLYSLRWISDHELYLKQENNLVFNAYGNSVFLNSTDFEGCHSINDYSISPD 110
Qy 61 GQFILLVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 120
Db 111 GQFILLVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSVRIITWTGKEDIYINGITDMWYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSVRIITWTGKEDIYINGITDMWYEEVFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFVNTDSLSSTVNTATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFVNTDSLSSTVNTATSIQITAP 290
Qy 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCILVARQHIEMST 300
Db 291 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCILVARQHIEMST 350

QY 301 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKOCTFITKGTWEVIGIEAL 360
 DB 351 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKOCTFITKGTWEVIGIEAL 410
 QY 361 TSDLYIISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCOYYSVSFSKEAKYQOL 420
 DB 411 TSDLYIISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCOYYSVSFSKEAKYQOL 470
 QY 421 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFIILNETKFWYQOMIL 480
 DB 471 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFIILNETKFWYQOMIL 530
 QY 481 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 540
 DB 531 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 590
 QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGVYTSWVLGSGGVKFCG 600
 DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGVYTSWVLGSGGVKFCG 650
 QY 601 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
 DB 651 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
 QY 661 VHFQSAQISKALVDGVDFQAMWYTDHGHASSTAHQHIYTHMSHFIKQCFSLP 716
 DB 711 VHFQSAQISKALVDGVDFQAMWYTDHGHASSTAHQHIYTHMSHFIKQCFSLP 766

RESULT 8

ADJ83981
 ID ADJ83981 standard; protein; 766 AA.

AC ADJ83981;

DT 06-MAY-2004 (first entry)

XX Human full-length colon dipeptidyl peptidase IV (DPPIV) protein.

XX Crystal; protein co-ordinate data; dipeptidyl peptidase IV; DPPIV;
 KW immunomodulatory; antidiabetic; antihypertensive; antitumor;
 KW antithyroid; antineoplastic; antidiabetic; antihypertensive; antitumor;
 KW immune response; diabetes; inflammation; multiple sclerosis;
 KW Grave's disease; chronic rheumatoid arthritis; AIDS; cancer; human;
 KW colon; enzyme.

XX Homo sapiens.

XX W02004011640-A1.

XX 05-FEB-2004.

XX 28-JUL-2003; 2003WO-JP009523.

XX 29-JUL-2002; 2002US-0398761P.

XX (TANA) TANABE SEIYAKU CO.

XX Hiramatsu H, Kyono K, Shima H, Sugiyama S;

XX WPI; 2004-156830/15.

XX N-PSDB; ADJ83980.

XX New crystal of dipeptidyl peptidase IV capable of analyzing its three-
 PT dimensional structure, useful for designing, identifying, evaluating or
 PT searching an effector of the dipeptidyl peptidase IV.

XX Claim 3; SEQ ID NO 2; 332pp; English.

XX The invention relates to a novel crystal of a dipeptidyl peptidase IV
 CC (DPPIV) which is sufficient to ensure a resolution capable of analysing
 CC its three-dimensional structure up to the side chain level by X-ray
 CC crystallographic structural analysis. The crystal of the invention

CC demonstrates immunomodulatory, antidiabetic, antinflammatory,
 CC neuroprotective, antithyroid, antineoplastic, antitumor, anti-HIV and
 CC cytoskeletal activities and may be useful for providing a three-dimensional
 CC structural coordinate as the information for designing, identifying, IV.
 CC The effector may be useful as a modulatory agent of immune response and
 CC as a therapeutic or prophylactic agent for diabetes, inflammation,
 CC multiple sclerosis, Grave's disease, chronic rheumatoid arthritis, AIDS
 CC or cancer. The current sequence is that of the human full-length colon
 CC dipeptidyl peptidase IV (DPPIV) protein of the invention.

XX Sequence 766 AA;

QY Query Match 100.0%; Score 3877; DB 8; Length 766;
 DB Best Local Similarity 100.0%; Pred. No. 0;
 Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTVRLKLYSLRWISDHELYLKQENNLVFNAYGNSVFLNSTDFFGHSINDYSISPD 60
 DB 51 NTVRLKLYSLRWISDHELYLKQENNLVFNAYGNSVFLNSTDFFGHSINDYSISPD 110
 QY 61 GQFILLLEYNVKQWRHSYTSYDIYDLNKKQLITEERIPNNTQWTVMSPVGHKLAYVYNN 120
 DB 111 GQFILLLEYNVKQWRHSYTSYDIYDLNKKQLITEERIPNNTQWTVMSPVGHKLAYVYNN 170
 QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSLMWSNPTFLAYAQFND 180
 DB 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSLMWSNPTFLAYAQFND 230
 QY 181 TEVPLIEYSPYDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSSTVNTATSIQITAP 240
 DB 231 TEVPLIEYSPYDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSSTVNTATSIQITAP 290
 QY 241 ASMLIGDHVLCVDTWATQERISLOWLRRIQNSVMDICDYDESSGRNCLVARQHIEMST 300
 DB 291 ASMLIGDHVLCVDTWATQERISLOWLRRIQNSVMDICDYDESSGRNCLVARQHIEMST 350
 QY 301 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKOCTFITKGTWEVIGIEAL 360
 DB 351 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKOCTFITKGTWEVIGIEAL 410
 QY 361 TSDLYIISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCOYYSVSFSKEAKYQOL 420
 DB 411 TSDLYIISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCOYYSVSFSKEAKYQOL 470
 QY 421 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFIILNETKFWYQOMIL 480
 DB 471 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKCLDFIILNETKFWYQOMIL 530
 QY 481 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 540
 DB 531 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 590
 QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGVYTSWVLGSGGVKFCG 600
 DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGVYTSWVLGSGGVKFCG 650
 QY 601 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
 DB 651 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
 QY 661 VHFQSAQISKALVDGVDFQAMWYTDHGHASSTAHQHIYTHMSHFIKQCFSLP 716
 DB 711 VHFQSAQISKALVDGVDFQAMWYTDHGHASSTAHQHIYTHMSHFIKQCFSLP 766

RESULT 9

ADJ75313
 ID ADJ75313 standard; protein; 766 AA.

XX AC ADJ75313;

XX DT 20-MAY-2004 (first entry)

XX	Marker gene related amino acid sequence SEQ ID NO:565.
XX	
XX	bronchial asthma; chronic obstructive pulmonary disease;
KW	respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW	gene therapy; marker.
XX	
OS	Homo sapiens.
XX	
EP	EP1394274-A2.
XX	
PD	03-MAR-2004.
XX	
XX	04-AUG-2003; 2003EP-00254857.
PF	
XX	06-AUG-2002; 2002JP-00229312.
PR	
XX	20-MAR-2003; 2003JP-00077212.
PR	
XX	(GENO-) GENOX RES INC.
PA	
XX	
PI	Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;
PI	WPI; 2004-193155/19.
DR	
XX	
PT	Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT	comparing the expression level of a marker gene in a biological sample
PT	from a subject with the expression level of the gene in a sample from a
PT	healthy subject.
XX	
XX	Example 11; SEQ ID NO 565; 24pp; English.
XX	
CC	The present invention describes a method of testing for bronchial asthma
CC	or chronic obstructive pulmonary disease. The method comprises
CC	determining the expression level of a marker gene in a biological sample
CC	from a subject, comparing the expression level determined with the
CC	expression level of the marker gene in a biological sample from a healthy
CC	subject, and judging whether the subject has bronchial asthma or chronic
CC	obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC	genes (S1) whose expression levels increase when respiratory epithelial
CC	cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC	whose expression levels decrease when respiratory epithelial cells are
CC	stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC	testing for bronchial asthma or chronic obstructive pulmonary disease;
CC	(2) a kit for screening for a candidate compound for a therapeutic agent
CC	to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC	an animal model for bronchial asthma or chronic obstructive pulmonary
CC	disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC	method for producing an animal model for bronchial asthma or chronic
CC	obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC	asthma or chronic obstructive pulmonary disease, comprising the compound,
CC	a marker gene or an antisense nucleic acid corresponding to a portion of
CC	the marker gene, a ribozyme, a polynucleotide that suppresses the
CC	expression of the gene through an RNAi effect or an antibody recognising
CC	a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC	bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC	probe has been immobilised to assay a marker gene. (I) has respiratory
CC	and antiasthmatic activities, and can be used in gene therapy. The method
CC	is useful for testing for or screening for a therapeutic agent for
CC	bronchial asthma or chronic obstructive pulmonary disease. The present
CC	sequence is used in the exemplification of the present invention.
XX	
SQ	Sequence 766 AA;
	Query Match 100.0%; Score 3877; DB 8; Length 766;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 NTYRLKLSLRWISDHELYLKQNNILVFNAEYGNSSVFLFNSTDFEFGHSINDYSISP 60
DB	51 NTYRLKLSLRWISDHELYLKQNNILVFNAEYGNSSVFLFNSTDFEFGHSINDYSISP 110
QY	61 GQFILLEVNYVKQWRHSYASDYIDLNKKQLITEERIPNNOTQVWTSPVGHKLAYVNN 120

XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX WPI; 2004-420067/39.
XX N-PSDB; ADO19397.
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthropathy.
XX Claim 7; SEQ ID NO 328; 1731pp; English.
XX The invention relates to human PRO polypeptides and the polynucleotides
XX encoding them. The polypeptides and polynucleotides are useful for
XX treating and diagnosing immune related disorders in mammals. The immune
XX related disorders include systemic lupus erythematosus, rheumatoid
XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic
XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
XX mellitus, immune-mediated renal disease, demyelinating diseases of the
XX central or peripheral nervous system, demyelinating polyneuropathy,
XX Guillain-Barre syndrome and chronic inflammatory demyelinating
XX polyneuropathy. This sequence represents a human PRO polypeptide of the
XX invention.
XX Sequence 766 AA;
SQ
Query Match 100.0%; Score 3877; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NTYRLKLYSLRWISDHVELYKQENNLVFNAYGSSVFLNSTDFFGHGSHINDYSISPD 60
DB |||||
QY 51 NTYRLKLYSLRWISDHVELYKQENNLVFNAYGSSVFLNSTDFFGHGSHINDYSISPD 110
DB |||||
QY 61 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 120
DB |||||
QY 111 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 170
DB |||||
QY 121 DIYKIEPNLPSYRITWTGKEDIYNGITDWYEEVFSAYSALWSPNGTFLAYAQFND 180
DB |||||
QY 171 DIYKIEPNLPSYRITWTGKEDIYNGITDWYEEVFSAYSALWSPNGTFLAYAQFND 230
DB |||||
QY 181 TEVPLIEYSFYSDLSQVPTKVRVPYKAGAVNPTKFFVNTDSLSSVTNATSIQITAP 240
DB |||||
QY 231 TEVPLIEYSFYSDLSQVPTKVRVPYKAGAVNPTKFFVNTDSLSSVTNATSIQITAP 290
DB |||||
QY 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 300
DB |||||
QY 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 350
DB |||||
QY 301 TGWVGRPRPSPHPTLDGNSFYKIIISNEBGRHICYQIDKKOCTFTTKGTWEVIGIEAL 360
DB |||||
QY 351 TGWVGRPRPSPHPTLDGNSFYKIIISNEBGRHICYQIDKKOCTFTTKGTWEVIGIEAL 410
DB |||||
QY 361 TSDLYLYISNEYKMGPCGRNLKYQLSDYTKVTCLSCELPERCOYVSFSKEAKYQL 420
DB |||||
QY 411 TSDLYLYISNEYKMGPCGRNLKYQLSDYTKVTCLSCELPERCOYVSFSKEAKYQL 470
DB |||||
QY 421 RCGSGPLPLTLHSSVNDKGLRVLEDNSALDKMLQNVMPKSKLDLFIILNETKFWYQML 480
DB |||||
QY 471 RCGSGPLPLTLHSSVNDKGLRVLEDNSALDKMLQNVMPKSKLDLFIILNETKFWYQML 530
DB |||||
QY 481 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 540
DB |||||
QY 531 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 590
DB |||||
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGYVTSWLVGSSGVFKCG 600
DB |||||
QY 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGYVTSWLVGSSGVFKCG 650
DB |||||

QY 601 IAVAPVSRWEYDSVYTERYMGVGLPTPEDNLHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB |||||
QY 651 IAVAPVSRWEYDSVYTERYMGVGLPTPEDNLHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
DB |||||
QY 661 VHFQOSQAISKALVDVGVDFQAMKYTDEDHGIASSTAHOHIYTHMSHFQKCFSLP 716
DB |||||
QY 711 VHFQOSQAISKALVDVGVDFQAMKYTDEDHGIASSTAHOHIYTHMSHFQKCFSLP 766
DB |||||
RESULT 11
ADO19806
ID ADO19806 standard; protein; 766 AA.
XX AC ADO19806;
XX DT 12-AUG-2004 (first entry)
XX DE Human PRO polypeptide #365.
XX KW Human; PRO; immune related disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
KW diabetes mellitus; renal disease; demyelinating disease;
KW central nervous system; peripheral nervous system;
KW demyelinating polyneuropathy; Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX OS Homo sapiens.
XX PN WO2004043361-A2.
XX PD 27-MAY-2004.
XX PF 06-NOV-2003; 2003WO-US035268.
XX PR 08-NOV-2002; 2002US-0425235P.
XX PA (GETH) GENENTECH INC.
XX PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX WPI; 2004-420067/39.
XX N-PSDB; ADO19805.
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthropathy.
XX Claim 7; SEQ ID NO 730; 1731pp; English.
XX The invention relates to human PRO polypeptides and the polynucleotides
XX encoding them. The polypeptides and polynucleotides are useful for
XX treating and diagnosing immune related disorders in mammals. The immune
XX related disorders include systemic lupus erythematosus, rheumatoid
XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic
XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
XX mellitus, immune-mediated renal disease, demyelinating diseases of the
XX central or peripheral nervous system, demyelinating polyneuropathy,
XX Guillain-Barre syndrome and chronic inflammatory demyelinating
XX polyneuropathy. This sequence represents a human PRO polypeptide of the
XX invention.
XX Sequence 766 AA;
SQ
Query Match 100.0%; Score 3877; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NTYRLKLYSLRWISDHVELYKQENNLVFNAYGSSVFLNSTDFFGHGSHINDYSISPD 60

Db 51 NTVRLKLYSLRWISDHEYLKQENNLVFNAYGNSVFLNSTDFEFGHSINDYSISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 290
Qy 241 ASMLIGDHYLCDVWTWATERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHYLCDVWTWATERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKDKCTFTTKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKDKCTFTTKGTWEVIGIEAL 410
Qy 361 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSPKLDFFILNETKFWYQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSPKLDFFILNETKFWYQMIL 530
Qy 481 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDFGRSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDFGRSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSMLVSGSGGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSMLVSGSGGVFKCG 650
Qy 601 IAVAPVSRWEYYSVYTERYMGILPTPEDNLNDRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYYSVYTERYMGILPTPEDNLNDRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHFIKQCFSLP 766

RESULT 12
ADO71612
ID ADO71612 standard; protein; 766 AA.

XX
AC ADO71612;
AC ADO71612;
DT 26-AUG-2004 (first entry)
XX Amino acid sequence of a human CD26 protein.
XX CD26; chemotherapeutic; radiotherapeutic; cancer; cell growth;
XX dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;
XX cell cycle arrest; tumour; tumour necrosis; immune response; human.
OS Homo sapiens.
XX
XX WO2004045497-A2.
XX
XX 03-JUN-2004.
XX
XX
PF 15-MAY-2003; 2003WO-US015499.
XX
XX 17-MAY-2002; 2002US-0381606P.
XX
XX (TEXA) UNIV TEXAS SYSTEM.

XX
PI Dang NH, Morimoto C;
XX
XX MPI; 2004-420511/39.
XX N-PSDB; ADO71611, ADO71613.
PT Use of a CD26 composition, and a chemotherapeutic and/or a
PT radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell
PT cycle arrest, killing a cancer cell, treating cancer, or inducing tumor
PT regression or tumor necrosis.
XX
XX Claim 23; Page 151-153; 182pp; English.
XX
CC The specification describes a CD26 composition which, in conjunction with
CC chemotherapeutic or radiotherapeutic agents, is used for the treatment
CC and prevention of cancers. Expression of CD26 enhances the sensitivity of
CC the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26
CC is a dipeptidyl peptidase IV (DPPIV). The chemotherapeutic agent is a
CC topoisomerase II inhibitor. The CD26 composition of the invention is
CC useful for inhibiting the growth of a cell, inducing cell cycle arrest in
CC a cell, killing a cancer cell, potentiating the effect of a
CC chemotherapeutic agent and/or a radiotherapeutic agent on a tumour cell,
CC inducing or enhancing apoptosis of a cancer cell, treating cancer, or
CC inducing tumour regression or tumour necrosis. The CD26 composition is
CC further useful for increasing topoisomerase II expression in a cell, for
CC activating an antigen-presenting cell, or for potentiating immune
CC responses of an animal. The present sequence represents a CD26 protein,
CC and is encoded by vectors which are used to produce compositions of the
CC invention.
XX
XX Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTVRLKLYSLRWISDHEYLKQENNLVFNAYGNSVFLNSTDFEFGHSINDYSISPD 60
Db 51 NTVRLKLYSLRWISDHEYLKQENNLVFNAYGNSVFLNSTDFEFGHSINDYSISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 290
Qy 241 ASMLIGDHYLCDVWTWATERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHYLCDVWTWATERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKDKCTFTTKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKDKCTFTTKGTWEVIGIEAL 410
Qy 361 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSPKLDFFILNETKFWYQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSPKLDFFILNETKFWYQMIL 530
Qy 481 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDFGRSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDFGRSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSMLVSGSGGVFKCG 600

Db 591 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMLVSGSGVFKCG 650
Qy 601 IAVAPVSRWEYDVSVTYRYMGLPTPDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVTYRYMGLPTPDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFQKCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFQKCFSLP 766

RESULT 13
AD071644
ID ADO71644 standard; protein; 766 AA.
XX
AC ADO71644;
XX
DT 26-AUG-2004 (first entry)
XX
DE Amino acid sequence of a human CD26 protein.
XX
KW CD26; chemotherapeutic; radiotherapeutic; cancer; cell growth;
KW dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;
KW cell cycle arrest; tumour; tumour necrosis; immune response; human.
XX
OS Homo sapiens.
XX
XN WO2004045497-A2.
XX
PD 03-JUN-2004.
XX
PF 15-MAY-2003; 2003WO-US015499.
XX
PR 17-MAY-2002; 2002US-0381606P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Dang NH, Morimoto C;
XX
DR WPI; 2004-420511/39.
DR N-PSDB; ADO71643.
XX
PT Use of a CD26 composition, and a chemotherapeutic and/or a
PT radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell
PT cycle arrest; killing a cancer cell, treating cancer, or inducing tumor
PT regression or tumor necrosis.
XX

PS Claim 23; Page 175-176; 182pp; English.
XX
XX The specification describes a CD26 composition which, in conjunction with
XX chemotherapeutic or radiotherapeutic agents, is used for the treatment
XX of cancer. Expression of CD26 enhances the sensitivity of
XX the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26
XX is a dipeptidyl peptidase IV (DPPIV). The chemotherapeutic agent is a
XX topoisomerase II inhibitor. The CD26 composition of the invention is
XX useful for inhibiting the growth of a cell, inducing cell cycle arrest in
XX a cell, killing a cancer cell, potentiating the effect of a
XX chemotherapeutic agent and/or a radiotherapeutic agent on a tumour cell,
XX inducing or enhancing apoptosis of a cancer cell, treating cancer, or
XX inducing tumour regression or tumour necrosis. The CD26 composition is
XX further useful for increasing topoisomerase II expression in a cell, for
XX activating an antigen-presenting cell, or for potentiating immune
XX responses of an animal. The present sequence represents a CD26 protein,
XX and is encoded by vectors which are used to produce compositions of the
XX invention.
SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTDFFGHISINDYSISPD 60
Db 51 NTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTDFFGHISINDYSISPD 110
Qy 61 GQFILLEYNVYKQWRHSYASYDIYDLNKRQLITEERI PNTQWTVTSPVGHKLAYYWN 120
Db 111 GQFILLEYNVYKQWRHSYASYDIYDLNKRQLITEERI PNTQWTVTSPVGHKLAYYWN 170
Qy 121 DIYVKLEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSAALWSPNGTFLAYAOFPND 180
Db 171 DIYVKLEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSAALWSPNGTFLAYAOFPND 230
Qy 181 TEVPLIEYSFYSDSLQYPKTVRPYPKAGAVNPTVKFFVYVNTDSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDSLQYPKTVRPYPKAGAVNPTVKFFVYVNTDSLSSVTNATSIQITAP 290
Qy 241 ASMLIGDHVLCVDTWATQBRISLOWLRRIONYSVMDICDYDESSGRWNCILVARQHIEMST 300
Db 291 ASMLIGDHVLCVDTWATQBRISLOWLRRIONYSVMDICDYDESSGRWNCILVARQHIEMST 350
Qy 301 TGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKDKCTFITKGTWVIGIEAL 360
Db 351 TGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKDKCTFITKGTWVIGIEAL 410
Qy 361 TSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 411 TSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFTILNETFPWYQML 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFTILNETFPWYQML 530
Qy 481 PPFPDKSKYPLLLDVYAGPCOKADTVFRLNWTYLASTENIIVASFDCRSGSYQGDKI 540
Db 531 PPFPDKSKYPLLLDVYAGPCOKADTVFRLNWTYLASTENIIVASFDCRSGSYQGDKI 590
Qy 541 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMLVSGSGVFKCG 600
Db 591 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMLVSGSGVFKCG 650
Qy 601 IAVAPVSRWEYDVSVTYRYMGLPTPDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVTYRYMGLPTPDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFQKCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFQKCFSLP 766

RESULT 14
ABM80355
ID ABM80355 standard; protein; 766 AA.
XX
AC ABM80355;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO80881, SEQ:895.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX

PP 29-SEP-2003; 2003WO-US028547.
 XX
 PR 02-OCT-2002; 2002US-0414971P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Wu TD, Zhang Z, Zhou Y;
 XX
 DR WPI; 2004-347921/32.
 DR N-PSDB; ACN37783.
 XX
 New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.
 XX
 Claim 12; SEQ ID NO 895; 7273pp; English.
 XX
 The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT polypeptide of the invention
 XX
 XX Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 8; Length 766;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NTRYLKLYSRWISDHELYLKQENNLVFNAYGNSVFLNSTDFEGHSINDYSISP 60
 DB 51 NTRYLKLYSRWISDHELYLKQENNLVFNAYGNSVFLNSTDFEGHSINDYSISP 110
 QY 61 GQFILLEYNVYKQWRHSYTASYDIYDLNKRQLITEERIPNNTQVWTSVPVGHKLAYVWN 120
 DB 111 GQFILLEYNVYKQWRHSYTASYDIYDLNKRQLITEERIPNNTQVWTSVPVGHKLAYVWN 170
 QY 121 DIVVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYSALWSPNGTFLAYAQFND 180
 DB 171 DIVVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYSALWSPNGTFLAYAQFND 230
 QY 181 TEVPLIEYSFYSDLESQPKTVRPVYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 240
 DB 231 TEVPLIEYSFYSDLESQPKTVRPVYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 290
 QY 241 ASMLIGHYICDVWTATQERISLOWLRRIQNYSVMDICDYDESSGRWNCVLVARQHIE MST 300
 DB 291 ASMLIGHYICDVWTATQERISLOWLRRIQNYSVMDICDYDESSGRWNCVLVARQHIE MST 350
 QY 301 TGVWGRFPSPHPFTLDGNSFYKLIISNEEGRVHCYFQIDKDKCTFTTKGTWEVIGIEAL 360
 DB 351 TGVWGRFPSPHPFTLDGNSFYKLIISNEEGRVHCYFQIDKDKCTFTTKGTWEVIGIEAL 410
 QY 361 TSDYLYYISNEYKMGPGGRNLKYQLSDYTKVTCLSCELPNRCQYVSFSKEAKYQL 420
 DB 411 TSDYLYYISNEYKMGPGGRNLKYQLSDYTKVTCLSCELPNRCQYVSFSKEAKYQL 470

QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFWQMIL 480
 DB 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFWQMIL 530
 QY 481 PPHFDSKKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDRGSGYQGDKI 540
 DB 531 PPHFDSKKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDRGSGYQGDKI 590
 QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIALIWGWSYGGYVTSWVLGSGSGVFKCG 600
 DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIALIWGWSYGGYVTSWVLGSGSGVFKCG 650
 QY 601 IAVAPVSRWEYVDSVYTERYMGILPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
 DB 651 IAVAPVSRWEYVDSVYTERYMGILPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
 QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFHKQCFSLP 716
 DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFHKQCFSLP 766
 RESULT 15
 ADP54458
 ID ADP54458 standard; protein; 766 AA.
 XX
 AC ADP54458;
 XX
 DT 18-NOV-2004 (first entry)
 DE
 XX Human PRO protein sequence SEQ ID NO:434.
 XX human; PRO; immune related disease; inflammatory immune response;
 KW immune response stimulation; anti-allergic; antianaemic; antiarthritic;
 KW antitachmic; antidiabetic; antiinflammatory; antipsoriatic;
 KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
 KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
 KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
 KW virucide; gene therapy.
 XX
 OS Homo sapiens.
 XX
 XX WO2004039956-A2.
 XX
 PD 13-MAY-2004.
 XX
 PF 28-OCT-2003; 2003WO-US034381.
 XX
 PR 29-OCT-2002; 2002US-0422472P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
 PI Wood WI, Wu TD;
 XX
 XX WPI: 2004-376182/35.
 DR N-PSDB; ADP54457.
 XX
 PT New PRO polynucleotides and polypeptides, useful in useful in diagnosing
 PT and treating an immune related disease, e.g. systemic lupus
 PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
 PT stimulating an immune response.
 XX
 XX Claim 1; SEQ ID NO 434; 3009pp; English.
 PS
 XX The present invention describes an isolated PRO nucleic acid (I). Also
 CC described: (1) a vector comprising (1); (2) a host cell comprising the
 CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an
 CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
 CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
 CC antibody which specifically binds to a polypeptide of (4); (7) a
 CC composition of matter comprising a polypeptide of (4), an agonist or
 CC antagonist of the polypeptide or an antibody that binds to the
 CC polypeptide in combination with a carrier; (8) an article of manufacture

Search completed: January 30, 2006, 14:22:35
Job time : 138 secs

comprising a container, a label on the container and a composition of matter of (7); (9) a method of treating an immune related disease in a mammal; (10) a method for determining the presence of a PRO polypeptide in a sample suspected of having the polypeptide; (11) a method of diagnosing an immune related disease or an inflammatory immune response in mammal; (12) a method of identifying a compound that inhibits or mimics the activity of or expression of a gene encoding a PRO polypeptide; and (13) a method of stimulating the immune response in a mammal. The PRO sequences have anti-allergic, antianaemic, antiarthritic, antiaesthetic, antidiabetic, anti-inflammatory, antipruritic, antirheumatic, antithyroid, CNS, dermatological, gastrointestinal, haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular, nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and virucide activities, and can be used in gene therapy. The nucleic acid (I) and the encoded polypeptides, compositions, kits and methods are useful in diagnosing and treating an immune related disease and in stimulating an immune response. The present sequence represents a human PRO protein from the present invention.

XX Sequence 766 AA;

Query Match	100.0%;	Score 3877;	DB 8;	Length 766;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 716;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	NTYRLKLSLRWISDHEYLKQENNLVFNAYGNSVFLNENSTDFEFGHSINDYSISPD	60	
Db	51	NTYRLKLSLRWISDHEYLKQENNLVFNAYGNSVFLNENSTDFEFGHSINDYSISPD	110	
Qy	61	GQFILLBYNKKWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN	120	
Db	111	GQFILLBYNKKWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN	170	
Qy	121	DIYVKIEPNLPSRITWTGKEDIITNGITDWTVEEVFSAYSAWSPNGTFLAYAQFND	180	
Db	171	DIYVKIEPNLPSRITWTGKEDIITNGITDWTVEEVFSAYSAWSPNGTFLAYAQFND	230	
Qy	181	TEVPLIEYSFSDLSLOYPKTRVPYPKAGAVNPTKFFVNTDLSLSSVTNATSIQITAP	240	
Db	231	TEVPLIEYSFSDLSLOYPKTRVPYPKAGAVNPTKFFVNTDLSLSSVTNATSIQITAP	290	
Qy	241	ASMLIGDHYICDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWNCCLVARQHIMST	300	
Db	291	ASMLIGDHYICDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWNCCLVARQHIMST	350	
Qy	301	TGWGFRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKCOCTFITKGTWEVIGIEAL	360	
Db	351	TGWGFRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKCOCTFITKGTWEVIGIEAL	410	
Qy	361	TSYLYYIISNEYKMGPGGRNLYKIQLSDYTKVCLSCELNPERCOYYSVSFSKEAKYYQL	420	
Db	411	TSYLYYIISNEYKMGPGGRNLYKIQLSDYTKVCLSCELNPERCOYYSVSFSKEAKYYQL	470	
Qy	421	RCGSGPLPLTYLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFTILNETKFWQMIL	480	
Db	471	RCGSGPLPLTYLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFTILNETKFWQMIL	530	
Qy	481	PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI	540	
Db	531	PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI	590	
Qy	541	MHAINRLRGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSMVLGSGGVFKCG	600	
Db	591	MHAINRLRGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSMVLGSGGVFKCG	650	
Qy	601	IAPVPSRWEYSDSVYTERYMGLETPEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN	660	
Db	651	IAPVPSRWEYSDSVYTERYMGLETPEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN	710	
Qy	661	VHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP	716	
Db	711	VHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP	766	

THIS PAGE BLANK (USP 16)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2006, 14:14:31 ; Search time 41 Seconds
(without alignments)
1680.273 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877
Sequence: 1 NTVRLKLYSLRWISDHEYL.....AHQHITYHSHFIKQCFSLP 716

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3871	99.8	766	1 CDHU26	dipeptidyl-peptida
2	3353.5	86.5	792	1 A39314	dipeptidyl-peptida
3	3336	86.0	760	1 S23752	dipeptidyl-peptida
4	1930.5	49.8	759	2 I38593	fibroblast activat
5	1117	28.8	803	2 I68600	dipeptidyl aminope
6	1117	28.8	865	2 I54331	dipeptidyl aminope
7	1094	28.2	803	2 A41793	dipeptidyl aminope
8	940	24.2	793	2 T41703	dipeptidyl aminope
9	908	23.4	818	1 A30107	dipeptidyl aminope
10	791.5	20.4	711	2 S66261	X-Pro dipeptidyl-p
11	745	19.2	829	2 T19514	hypothetical prote
12	726	18.7	931	2 A49737	dipeptidyl aminope
13	720	18.6	779	2 T25173	hypothetical prote
14	716	18.5	799	2 T25173	hypothetical prote
15	628	16.2	738	2 A87516	dipeptidyl-peptida
16	594.5	15.3	741	2 JC5142	X-Pro dipeptidyl-p
17	590.5	15.2	743	2 T37700	probable dipeptidyl
18	445.5	11.5	931	2 T32919	hypothetical prote
19	332	8.6	795	2 F82858	dipeptidyl-peptida
20	272	7.0	657	2 E70025	probable acylamino
21	244.5	6.3	762	2 JC8016	acylaminoacyl-pept
22	222.5	5.7	622	2 F71174	hypothetical prote
23	221	5.7	591	2 H72474	probable acylamino
24	219	5.6	642	2 C71137	hypothetical prote
25	206.5	5.3	632	2 E75057	peptidase PAB1418
26	205	5.3	631	2 H75007	probable acylamino
27	205	5.3	683	2 E87495	prolyl oligopeptid
28	191	4.9	709	2 B82580	alanyl dipeptidyl
29	188	4.8	659	2 F72568	probable acylamino

30	183.5	4.7	721	2 T09631	probable acylamino
31	175	4.5	642	2 D87379	prolyl oligopeptid
32	174	4.5	676	2 C97775	acylamino-acid-rel
33	173.5	4.5	819	2 B87580	conserved hypothet
34	168	4.3	618	2 D87651	prolyl oligopeptid
35	163.5	4.2	761	2 S44807	F44B9.1 protein -
36	156.5	4.0	674	2 B84381	acylaminoacyl-pept
37	154	4.0	637	2 S75772	hypothetical prote
38	153.5	4.0	606	2 T35378	probable peptide h
39	149	3.8	582	2 D72636	probable acylamino
40	147.5	3.8	572	2 F72455	probable acylamino
41	147	3.8	598	2 F84199	hypothetical prote
42	146.5	3.8	732	1 S07624	acylaminoacyl-pept
43	145.5	3.8	683	2 AB0217	oligopeptidase B (
44	143	3.7	536	2 F90299	acylaminoacyl-pept
45	142.5	3.7	745	2 T33751	hypothetical prote

ALIGNMENTS

RESULT 1

CDHU26

dipeptidyl-peptidase IV (EC 3.4.14.5) - human

N;Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence revision 23-Aug-1996 #text change 09-Jul-2004

C;Accession: S24313, B42408, A42408, B61136, S59510, I56154, S59857, S15520

R;Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikehard, Y.

Biochim. Biophys. Acta 1131, 333-336, 1992

A;Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a se

A;Reference number: S24313, MUID:92329551, PMID:1352704

A;Accession: S24313

A;Molecule type: mRNA

A;Residues: 1-6, 'I', 8-766 <MIS>

A;Cross-references: UNIPROT:P27487; UNIPARC:UPI000016AE80; EMBL:X60708; NID:G35335; PID:

R;Barmaul, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, A.

J. Biol. Chem. 267, 4824-4833, 1992

A;Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cance

IV mRNA levels during cell differentiation.

A;Reference number: A42408, MUID:92165847, PMID:1347043

A;Accession: B42408

A;Molecule type: mRNA

A;Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-766 <DAR1>

A;Cross-references: UNIPARC:UPI0000052ACB; GB:M80536; NID:G181569; PIDN:AAA52308.1; PID

A;Experimental source: intestine

A;Note: this sequence corresponds with the author's translation

A;Accession: A42408

A;Molecule type: mRNA

A;Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-711, 'G', 713-766 <DAR2>

A;Cross-references: UNIPARC:UPI0000172A2B; GB:M80536; NID:G181569; PIDN:AAA52308.1; PID

A;Note: sequence extracted from NCBI backbone (NCBI:83986, NCBI:P.83988); this sequence

R;Gorvel, J.P.; Ferrero, A.; Chambrault, L.; Rigal, A.; Bonicel, J.; Maroux, S.

Gastroenterology 101, 618-625, 1991

A;Title: Expression of sucrose-isomaltase and dipeptidylpeptidase IV in human small int

A;Reference number: A61136; MUID:91317403; PMID:1677636

A;Accession: B61136

A;Molecule type: protein

A;Residues: 1-15, 'X', 17-22 <GOR>

A;Cross-references: UNIPARC:UPI0000172A2C

R;Boehm, S.K.; Gum Jr., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S.

Biochem. J. 311, 835-843, 1995

A;Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a

A;Reference number: S59510; MUID:96067599; PMID:7487939

A;Accession: S59510

A;Molecule type: DNA

A;Residues: 1-31 <BOE>

A;Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:G1195574; PIDN:AAB35614.1; PI

R;Tanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg,

J. Immunol. 149, 481-486, 1992

A;Title: Cloning and functional expression of the T cell activation antigen CD26.

A;Reference number: I56154; MUID:92325476; PMID:1352530

A;Accession: I56154

A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-436, 'S', 438-766 <TAN>
A:Cross-references: UNIPARC:UPI000004F7BF; GB:W74777; NID:G180082; PIDN:AAA51943.1; PID: R;Abbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W.
Immunogenetics 40, 331-338, 1994
A:Title: Genomic organization, exact localization, and tissue expression of the human CD
A:Reference number: S59857; MUID:95012454; PMID:7927537
A:Accession: S59857
A:Molecule type: DNA
A:Residues: 1-436, 'S', 438-766 <ABB>
A:Cross-references: UNIPARC:UPI000004F7BF; EMBL:UL13734
C:Genetics:
A:Gene: GDB:DPP4
A:Cross-references: GDB:1125239; OMIM:102720
A:Map position: 2q24.3-2q24.3
A:Introns: 2/3; 32/1; 65/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/3
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; homodimer; proteinase; transmembr
F:1-6/Domain: intracellular #status predicted <INT>
F:7-28/Domain: transmembrane #status predicted <TN>
F:29-766/Domain: extracellular #status predicted <EXT>
F:85,92,150,219,229,281,321,520,685/Binding site: carbohydrate (Asn) (covalent) #status
F:630,708,740/Active site: Ser, Asp, His #status predicted

Query Match 99.8%; Score 3871; DB 1; Length 766;
Best Local Similarity 99.9%; Pred. No. 5.9e-260;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTVRLKLYSRWTSDEHYLKQENNLVFNFAEYGNSSVFLENSTFBFGHSINDYISPD 60
DB 51 NTVRLKLYSRWTSDEHYLKQENNLVFNFAEYGNSSVFLENSTFBFGHSINDYISPD 110
QY 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKLQGLITEERIPNNTQVMTSPVGHKLAIVVNN 120
DB 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKLQGLITEERIPNNTQVMTSPVGHKLAIVVNN 170
QY 121 DIYVKIEPNLPSRIWTGKEDIYNGITDWTVEEVEFSAISALWSPNGTFLAYAQFND 180
DB 171 DIYVKIEPNLPSRIWTGKEDIYNGITDWTVEEVEFSAISALWSPNGTFLAYAQFND 230
QY 181 TEVPLEIYSYSDSLQYPKTVRPVYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
DB 231 TEVPLEIYSYSDSLQYPKTVRPVYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290
QY 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIEMST 300
DB 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIEMST 350
QY 301 TGVGGRFRPSEPHFTLDGNSFYKIIISNEGYRHICYFOIDKKDCTFTTKGTWEVIGIEAL 360
DB 351 TGVGGRFRPSEPHFTLDGNSFYKIIISNEGYRHICYFOIDKKDCTFTTKGTWEVIGIEAL 410
QY 361 TSDLYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSFVSKEAKYYQL 420
DB 411 TSDLYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSFVSKEAKYYQL 470
QY 421 RCGSGPLLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSEKGLDFIILNETKFWYQML 480
DB 471 RCGSGPLLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSEKGLDFIILNETKFWYQML 530
QY 481 PPHFDKSKYPLLLDLYAGPCSKADTVFRLNWKATYLASTENIIVASFGRSGYQGDKI 540
DB 531 PPHFDKSKYPLLLDLYAGPCSKADTVFRLNWKATYLASTENIIVASFGRSGYQGDKI 590
QY 541 MHAINRLRGTFEVEDQIEAARQPSKMGFVDNKRKIALWGSYGGYVTSMLVSGSGVFKCG 600
DB 591 MHAINRLRGTFEVEDQIEAARQPSKMGFVDNKRKIALWGSYGGYVTSMLVSGSGVFKCG 650
QY 601 IAVAPVSRWEYDSVYTERYMGUPTPEDNLDRHNRSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPVSRWEYDSVYTERYMGUPTPEDNLDRHNRSTVMSRAENFKQVEYLLIHGTADDN 710

QY 661 VHFQSOAQISKALVDGVDFQAMWYTDDEHGJASSTAHOHIYTHMSHFIRKQCFSLP 716
DB 711 VHFQSOAQISKALVDGVDFQAMWYTDDEHGJASSTAHOHIYTHMSHFIRKQCFSLP 766

RESULT 2
A39914 dipeptidyl-peptidase IV (EC 3.4.14.5), membrane-bound form precursor - rat
N;Alternate names: GP110; membrane glycoprotein 110K; OX-61
N;Contains: dipeptidyl-peptidase IV, soluble form
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C:Accession: A39914; A33315; B33315; A60730; A42203; S38949; A31781
R;Hong, W.; Doyle, D.
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987
A:Title: cDNA cloning for a bile canaliculus domain-specific membrane glycoprotein of r
A:Reference number: A39914; MUID:88068516; PMID:3479775
A:Accession: A39914
A:Molecule type: mRNA
A:Residues: 1-792 <HON>
A:Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN
R;Ogata, S.; Miuma, Y.; Ikehara, Y.
J. Biol. Chem. 264, 3596-3601, 1989
A:Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA a
A:Reference number: A33315; MUID:89123496; PMID:2563382
A:Accession: A33315
A:Molecule type: mRNA
A:Residues: 1-37, 'A', 39-182, 'I', 184-331, 'T', 333-351, 'C', 353-393, 'V', 395-561, 'L', 563-623
A:Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:G203973; PIDN:AAA41096.1; PID
A:Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332
A:Accession: B33315
A:Molecule type: protein
A:Residues: 1-20;35-54;427-443;505-509;511-520;530-538;593-600;602-608;618-627 <HO2>
A:Cross-references: UNIPARC:UPI0000172A2E; UNIPARC:UPI0000172A30; UNIPARC:UPI0000172A31
A36; UNIPARC:UPI0000172A37
R;McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.
Hepatology 11, 534-544, 1990
A:Title: Identification of the bile canalicular cell surface molecule GP110 as the ecto
quence.
A:Reference number: A60730; MUID:90228896; PMID:1970322
A:Accession: A60730
A:Molecule type: protein
A:Residues: 28-47, 'XX', 50-53 55-58 <MCC>
A:Cross-references: UNIPARC:UPI0000172A38
R;Ogata, S.; Miuma, Y.; Tsuji, E.; Takami, N.; Oda, K.; Ikehara, Y.
Biochemistry 31, 2582-2587, 1992
A:Title: Identification of the active site residues in dipeptidyl peptidase IV by affi
A:Reference number: A42203; MUID:92190188; PMID:1347701
A:Accession: A42203
A:Molecule type: protein
A:Residues: 'R', 625-630, 'X', 632-648 <OG2>
A:Cross-references: UNIPARC:UPI0000172A39
R;Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993
A:Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl
A:Reference number: S38949; MUID:94128239; PMID:7905271
A:Accession: S38949
A>Status: preliminary
A:Molecule type: protein
A:Residues: 281-302 <IWA>
A:Cross-references: UNIPARC:UPI0000172A3A
R;Hong, W.; Doyle, D.
J. Biol. Chem. 263, 16892-16898, 1988
A:Title: Membrane orientation of rat gp110 as studied by in vitro translation.
A:Reference number: A31781; MUID:89034185; PMID:3182821
A:Accession: A31781
A:Molecule type: mRNA
A:Residues: 1-40 <HO3>
A:Cross-references: UNIPARC:UPI0000172A3B
A:Comment: this protein is localized to the bile canaliculus, which is the apical domai
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; homodimer; liver; serine protein
F:1-792/Product: dipeptidyl-peptidase, membrane-bound form #status experimental <MATM>

QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMLVSGSGGVFKCG 600
DB 585 MHAINRRLGTFEVEDQIEAARQFVKMGFVDKRVAVWGSYGGYVTSMLVSGSGGVFKCG 644
QY 601 IAVAPSRWEYDSVYTERYMGVLPEDNLDHYRNSVMSRAENFKQVEYLLIHGTADDN 660
DB 645 IAVAPSRWEYDSVYTERYMGVLPEDNLDHYRNSVMSRAENFKQVEYLLIHGTADDN 704
QY 661 VHFQSOAQISKALVDVDFQAMWYTDDEHGIIASSTAHQIYTHMSHFQKQCSFL 715
DB 705 VHFQSOAQISKALVDVDFQAMWYTDDEHGIIASSTAHQIYTHMSHFQKQCSFL 759

RESULT 4
I38593
fibroblast activation protein-alpha - human
N;Alternate names: FAP-alpha
C;Species: Homo sapiens (man)
C;Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 29-Aug-1997
C;Accession: I38593
R;Scanlan, M.J.; Raj, B.; Calvo, B.; Garin-Chesa, P.; Sanz-Moncasi, M.P.; Healey, J.; O'Brien, J.; et al. Proc. Natl. Acad. Sci. U.S.A. 91, 5657-5661, 1994
A;Title: Molecular cloning of fibroblast activation protein alpha, a member of the serin
A;Reference number: I38593; MUID:94261645; PMID:7911242
A;Accession: I38593
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-759 <RES>
A;Cross-references: UNIPARC:UPI000175B7D; EMBL:U09278; NID:g507749; PID:g507750
C;Genetics:
C;Superfamily: transmembrane protein

Query Match 49.8%; Score 1930.5; DB 2; Length 759;
Best Local Similarity 49.2%; Pred. No. 13e-125;
Matches 388; Conservative 134; Mismatches 201; Indels 35; Gaps 11;

QY 2 TYRLKLYSLRWISDHELYLKQ-ENNILVFNAEYGNSSVFLENSTFDFGHSIN--DYSIS 58
DB 51 TFSYKTFPFWNISGQBYLHQSADNNVLYNIETGQSYTILSNRTM---KSNVNASNYGLS 106
QY 59 PDGQFILLEYNVQWRHSYASDIYDLNKRQLITEERIPNNTQWVTSVPGHKLAYTW 118
DB 107 PDRQFYLESQYSLKWRYSYATYTYIYDLSNGEFVRGNELPRPIQYLCSPVGSKLAYVY 166
QY 119 NNDIYVKBPNLPSYRITWTGKEDIYNGITDMVYEEVFSAYSAALWSPNGTFLAYAQF 178
DB 167 QNNIYILKQRPDPPQITFNGRENKIPNGIPDWVYEEEMLPKYALWSPNGKFLAYAEF 226
QY 179 NDEVEPLIERYSFYSDLESQYPKTVRPYPKAGAVNPTVKFFVNTDLSSTVNTATSQIT 238
DB 227 NDKDIPVIAYSYIGDE--QYPTINIPYKAGAKNPVVRIFIIDTTYPAVVGQ---EVP 281
QY 239 APASMLIGHYLCVDTWATQERISLOWLRRIQNYVMDICDYDESSGRWNCVLVAROHIE 298
DB 282 VPAMIASDYSYFSLWLTWTDERYCLOQLKRVQNVSVLSICDFREDQWDCPKTOHIEE 341
QY 299 STTGWGRFRPSPHFTLDGNSFYKLIISNEEGYRHICYFIQDKKQCTFTTKGTWEVIGIE 358
DB 342 SRTGWAGGFVSRPVFSYDAISYIKLPSDKDGKFKLHYIKOTVENAIQITSGKEAINTF 401
QY 359 ALTSYLYYISNEYKMGPGGRNLYKIQLSDY-TKVTCLSCELPNPERCQYYSVFSKEAKY 417
DB 402 RVTQDSLFSYSSNEFEYPPGRNRIYISGYPPSKCVCTHLRKERCQIYTSFSDYAKY 461
QY 418 YQLRCSGPGPLPYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSKKLDFFIILNETKFTWQ 477
DB 462 YALVCYGPPISTLHDGRTDQBIKLEENKELENALKNIQLPKESIKKLEVDITLWKY 521
QY 478 MILPPHFDKSKYPYLLLDVYAGCSQKADTVFRLNWTATYLASTENIIIVASFDRGSGYQG 537

DB 522 MILPQFDSKYPYLLIQVYGPCSQSVRSVFAVNWISYLASKEGVIALVDGRGTAFQG 581
QY 538 DKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGGYVTSMLVSGSGGVFKCG 587
DB 582 DKLLAVYRKLGVWEVEDQITAVRKFIEMGFIDEKRIAIWGSYEIFRITGPIPCINWWSFQ 641
QY 588 VYLGSGSGVFKGCIAPVPSRWYDYVYTERYMGVLPEDNLDHYRNSVMSRAENFKQ 647
DB 642 MWYSSGSL-QGLIRVLCIHR-----BIHSPNKDDNLEHYKNSVMSRAEYFRN 690
QY 648 VEYLLIHGTADDNVHFFQSOAQISKALVDVDFQAMWYTDDEHGIIASSTAHQIYTHMSH 707
DB 691 VDYLLIHGTADDNVHFFQSOAQISKALVNAQVDFQAMWYSDQNHL-SGLSTNHLTYHMT 749
QY 708 FIKQCFSL 715
DB 750 FLKQCFSL 757

RESULT 5
I68600
dipeptidyl aminopeptidase like protein - human
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C;Accession: I68600
R;Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K. Hum. Mol. Genet. 2, 1037-1039, 1993
A;Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-relat
A;Reference number: I54331; MUID:93372805; PMID:8103397
A;Accession: I68600
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-803 <RES>
A;Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:G306707; PIDN:AAA35761.1; PID
C;Superfamily: dipeptidyl-peptidase IV

Query Match 28.8%; Score 1117; DB 2; Length 803;
Best Local Similarity 33.2%; Pred. No. 3e-69;
Matches 241; Conservative 150; Mismatches 290; Indels 44; Gaps 20;

QY 11 RWISDHELYLK-QENNILVFNAEYGNSSVFLENSTFDFGHSINDYISPDGQFILLLEYN 69
DB 87 KWISTEFTIYRQKGTVRLNVTWTSVLIEGKKIESL-RAIR-YESIPREYALFSYN 144
QY 70 YVQWRHSYASDIYDL--NKRQLITEERIPN-NTQWVTSVPGHKLAYVWNNDIYVKI 126
DB 145 VEPIYQHSYTYGYVLSKIPHGDPQSLDPPPEVSNAKLQYAGWPKGQQLIFIFENNIYYCA 204
QY 127 EBNLPSYRITWTGKEDIYNGITDMVYEEVFSAYSAALWSPNGTFLAYAQNDTEVPLI 186
DB 205 HVGKQAIRVSTGKEGVYINGLSDWLYBEEILKTHIAHWSPDGTRLAYAANDSRVPI 264
QY 187 EYSFYSDESLOQYPKTVRPYPKAGAVNPTVKFFVNTDLSSTVNTATSQITAPASMLIG 246
DB 265 ELPTVTSIGI--YPTVKPYHPKAGSENPSISLHVI---GLNGPTH--DLEWMPDPPRKR 317
QY 247 DHYLCVDTWATQERISLOWLRRIQNYVMDICDYDESSGRWNCVLVAROHIEHSTTGWGR 306
DB 318 EYIITWVKWATSKVAVTWNRAQNVSIITLC--DATTG----VCTKKH-EDESEAWLHR 370
QY 307 FRPSPHFTLDGNSFYKLIISNEEGYRHICYF-----QIDKDKCTFTTKGTWEVIGIE 358
DB 371 -QNEEFVSKDGRKFEFFIRAIPTQGRGKPYHITVSSSQPNSSNDNIQISTGDMVDVKIL 429
QY 359 AL--TSYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCTVTCLSCELPNPERCQYYSVFSKEAK 416
DB 430 AYDEKGNKIYFLSTE--DLPRRRQLYSANTEGNFRQCLSCDL-VENCITYFASFSHMD 486
QY 417 YQLRCSGPGPLPYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSKKLDFFIILNETKFTWY 476
DB 487 FELLKCEGPGVPMVTVHNTDKKKFQDLETNEHVKKAINDRQMPKVEYRDIIDDYNLPM 546

QY 477 QMILPHPHDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGGYQ 536
 Db 547 QILKPATFTDTTHYPLLLVVDGTPGSSQVAEKFEVSWETVMVSHGAVVVKCDGRSGGFQ 606
 QY 537 GDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIAGWSYGYVTSMWL---GSG 593
 Db 607 GTKLLHEVRRLGLLEKQDQMEAVRMLKEQYIDRTVAVFGKDYGGYLSTYILPAKGEN 666
 QY 594 SG-VFKCGIAPVPSRWEYVDSVYTERYMGLPPTPEDNLDH--YRNSVTMSRAENFKQVEY 650
 Db 667 QGQFTTCGSALSPTDPKLYASAFSERYLGL----HGLDNRAVEMTKVAHRVSALBEQQF 722
 QY 651 LLIHGTADDNVHQQSAQISKALVDVGVDFQAMWYTTDEHGIASSTAHQHIYTHMSHF 710
 Db 723 LIHTPTADEKIHQHTAELITQLIRKANYSLOIYPDESHYFTSSSLKQHLYSIINFFV 782
 QY 711 QCFSL 715
 Db 783 ECFRI 787

RESULT 6

I54331
 dipeptidyl aminopeptidase like protein - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C:Accession: I54331
 R:Yokotani, N.; Doi, K.; Wenthhold, R.J.; Wada, K.
 Hum. Mol. Genet. 2, 1037-1039, 1993
 A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
 A:Reference number: I54331; MUID:93372805; PMID:8103397
 A:Accession: I54331
 A:Status: preliminary;
 A:Molecule type: mRNA
 A:Residues: 1-865 <RES>
 A:Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:M96859; NID:g306705; PIDN:
 C:Superfamily: dipeptidyl-peptidase IV

Query Match 28.8%; Score 1117; DB 2; Length 865;
 Best Local Similarity 33.2%; Pred. No. 3.3e-69;
 Matches 241; Conservative 150; Mismatches 290; Indels 44; Gaps 20;

QY 11 RWISDHEYLK-QENNILVFNAYGNSSVFLENSTDFEGHSINDYSISPDGQFILLEYN 69
 Db 149 KWISDTEFYREQGTGTVRLNVNVTNTSTVLEGGKIESL-RAIR-YEISPDREYALFSYN 206
 QY 70 YVQWRHSYTSVDIYDL--NKRQLITEERIPN-NTQWTVTSPVGHKLAVVWNDIYVKI 126
 Db 207 VEPYQHSYTYGYVLSKIPHGDPQSLDPPPEVSNAKLQYAGMGPKGOQLIFIFENNIYCA 266
 QY 127 EPNLPSYRITWTKGEDIYNGITDWYEEVEFVSAYSALMWSPNGTFLAYAOFNDETVPLI 186
 Db 267 HVCKQAIRVSVTSGEGVIYNGLSDWLYEEELKTHAHWSPDGTFLAYAINDSRVPIM 326
 QY 187 EYSFYDESLSQYPKTVRVPYKAGAVNPVKFFVNTDLSSTVNTATSQITAPASMLIG 246
 Db 327 ELPTYTGSII--YPTVKPYHPKAGSNPISLHVI---GLNGPTH--DLEMPDPDPRMR 379
 QY 247 DHYLCQVWATQERISLOWLRRIQNVSMDCIDYDESSGRWNCVLARQHIEMSTTCGWVGR 306
 Db 380 EYITWVKWATSKVAVTLNRAQNVSIITLC--DATTG---VCTKKH-EDSEAWLHR 432
 QY 307 FRPSEPHFTLDGNSFYKIIISNEEGYRHICYF-----QIDKKDCTFTTKGTWEVIGIE 358
 Db 433 -QNEEPVFSKDGKRFFFRAIPQGGKGFPHITVSSSQPNSSNDNIQSITSGDWDVTKIL 491
 QY 359 AL--TSDLYYIINEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAK 416
 Db 492 AYDEKGNKIYFLSTE--DLPRRQLYSANTEGNFNQCLSCDL- VENTCYFVSASFHSD 548
 QY 417 YYQLRCGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLLFIILNETKFWY 476
 Db 549 FFLKCEGPGVPMVTVHNTTDDKKMFDELNETNEHVKAINDRQMPKVEYRDIEIDYDNLPM 608

QY 477 QMILPHPHDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGGYQ 536
 Db 609 QILKPATFTDTTHYPLLLVVDGTPGSSQVAEKFEVSWETVMVSHGAVVVKCDGRSGGFQ 668
 QY 537 GDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIAGWSYGYVTSMWL---GSG 593
 Db 669 GTKLLHEVRRLGLLEKQDQMEAVRMLKEQYIDRTVAVFGKDYGGYLSTYILPAKGEN 728
 QY 594 SG-VFKCGIAPVPSRWEYVDSVYTERYMGLPPTPEDNLDH--YRNSVTMSRAENFKQVEY 650
 Db 729 QGQFTTCGSALSPTDPKLYASAFSERYLGL----HGLDNRAVEMTKVAHRVSALBEQQF 784
 QY 651 LLIHGTADDNVHQQSAQISKALVDVGVDFQAMWYTTDEHGIASSTAHQHIYTHMSHF 710
 Db 785 LIHTPTADEKIHQHTAELITQLIRKANYSLOIYPDESHYFTSSSLKQHLYSIINFFV 844
 QY 711 QCFSL 715
 Db 845 ECFRI 849

RESULT 7

A41793
 dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A41793
 R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthhold, R.J.; Shimasaki, S.
 Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
 A:Title: Differential expression of two distinct forms of mRNA encoding members of a di-
 A:Reference number: A41793; MUID:92108018; PMID:1729689
 A:Accession: A41793
 A:Status: preliminary;
 A:Molecule type: mRNA
 A:Residues: 1-803 <WAD>
 A:Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A83C; GB:M76429; NID:g408719; PIDN:
 C:Superfamily: dipeptidyl-peptidase IV
 C:Key words: dipeptidylpeptide hydrolase; glycoprotein
 F:557,342/Binding site: carbohydrate (Aen) #status predicted

Query Match 28.2%; Score 1094; DB 2; Length 803;
 Best Local Similarity 32.9%; Pred. No. 1.2e-67;
 Matches 239; Conservative 153; Mismatches 287; Indels 48; Gaps 22;

QY 11 RWISDHEYLK-QENNILVFNAYGNSSVFLENSTDFEGHSINDYSISPDGQFILLEYN 69
 Db 87 KWISDKEFTYREQGKSVILRNVTNTSTVLEGGKIESL-RAIR-YEISPDREYALFSYN 144
 QY 70 YVQWRHSYTSVDIYDL--NKRQLITEERIPN-NTQWTVTSPVGHKLAVVWNDIYVKI 126
 Db 145 VEPYQHSYTYGYVLSKIPHGDPQSLDPPPEVSNAKLQYAGMGPKGOQLIFIFENNIYCA 204
 QY 127 EPNLPSYRITWTKGEDIYNGITDWYEEVEFVSAYSALMWSPNGTFLAYAOFNDETVPLI 186
 Db 205 HVCKQAIRVSVTSGEGVIYNGLSDWLYEEELKTHAHWSPDGTFLAYATINDSRVPM 264
 QY 187 EYSFYDESLSQYPKTVRVPYKAGAVNPVKFFVNTDLSSTVNTATSQITAPASMLIG 246
 Db 265 ELPTYTGS--VYPTAKPYHPKAGCENPISLHVI---GLNGPTH--DLEMPDPDPRMR 317
 QY 247 DHYLCQVWATQERISLOWLRRIQNVSMDCIDYDESSGRWNCVLARQHIEMSTTCGWVGR 306
 Db 318 EYITWVKWATSKVAVNLNRAQNVSIITLC--DATTG---VCTKKH-EDSEAWLHR 370
 QY 307 FRPSEPHFTLDGNSFYKIIISNEEG---YRHICYF-----QIDKKDCTFTTKGTWEVIGIE 358
 Db 371 -QNEEPVFSKDGKRFFFRAIPQGGKGFPHITVSSSQPNSSNDNIQSITSGDWDVTKI- 428
 QY 359 ALTSD-----YLYYIINEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKE 414
 Db 429 -LSYDEKRSQIYFLSTE--DLPRRQLYSASTVGSFNRQCLSCDL- VDNCTYFVSASFSPG 484

Db 136 NGQNLTVESITASPLKRLKLLIRTSVQNRHSTGSGVFVDKSSSF---EEIGNEVALA 192
Qy 106 TSPVGHKLAYVWNNDIYVIEPNLPYRITWTGKEDIINGIDWVYEEVFSAYSALW 165
Db 193 IWSPNNDIAVQDNNIYISAIKKTIRAVTNDGSSFLNGKFDWYEEVEFDDKAAW 252
Qy 166 WSPNGTFLAYQFNDTEVP--LIBYSYDESLOYKTVRPYPKAGAVNPTVKFFVNT 223
Db 253 WSPGTDYLAFLKIDSEGEFIIP-YVQDEKDIYPEMRISIKPKSGTPNPABELWY-- 309
Qy 224 DLSSTVNATSIQITAPASMLIGHYLCDDVTWATQERISLQWLRIQNSYVMDICDYDES 283
Db 310 ----SMKDCGTFPHRISGNKDGSLLTETVWVNGVNLVKTDRSSDILTVFLDITIAK 365
Qy 284 SGRWNCILVARQHIMSTTGM-----VGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQ 338
Db 366 TSN-----VVRN-----ESSNGMWETHTNLTFIIPANETFORPHGVVDILP-IGGYNHLAYFE 418
Qy 339 -IDKDCOTFITKGTWEVI-GIEALTS--DYLYISNEYKMGPGGRNLYKIQL----SDYTK 391
Db 419 NSNSSHYKTLTEGKWEVNGVFLAPDSMENRLYFISTRSKSTE--RHVYIIDLRSPNEIE 476
Qy 392 VTCLSCELNPERCOYYSVSFSEAKYVQLRCSGGLP---LYTLHS-----SYNDK 439
Db 477 VTDTSB-----GVYDVSFSRRRFGLLTYKGPVYQKIVDPHRSKAEKDKGNVLGK 530
Qy 440 GLRVLENSALDKMLQVQVPSKKLDFIILNETKFWYQ-----ILPPHFDK--SKYP 491
Db 531 SLYHLEKNEVLTKILEDYAVPRKSFRELNLGKDFGKDLVNSVEILLPNDPDETLSDHYP 590
Qy 492 LLDVYAGCSQKADTVFRLNWTATYLASTENIIVASPDGRSGVQGDKIMHAINRRLGTF 551
Db 591 VFFPAYGPNSSQVVKTFVSGFNEVWASQLNAIVVVDGRGTGKQDPRSLVRDLGDY 650
Qy 552 EVEDOIEAARQFSKMGVDNKRITAIWGSYGGYVTSWVLGSGSG-VFKGCIAPVSRWE 610
Db 651 EARDQISAAISLYSLTVDPKISLFGWSYGGYVTLTKLEKGGGRHFYKMGSAVPTDWR 710
Qy 611 YDVSVYTERYMGLTPTEDNLDHY-----RNSVTMSRAENFKQVEYLLIHGTAADNVHFQ 665
Db 711 FYDSVYTERY--HTQENFDGVSVSVHNTALQANRF-----LLMHGTGDDNVHFPQ 763
Qy 666 SAQISKALVDVGDV-FOAMWYTTDEDHGIIASSTAHOHIYTHMSHFIKQCF 713
Db 764 SLKFLDLLNGVENVDVHVPFSDHSIRVHNANVIVFDKLDWAKRAF 812

RESULT 10
S66261
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
C;Species: Flavobacterium meningosepticum
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S66261
R;Kobashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A;Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV gene from *P.*
A;Reference number: S66261; MUID:95314307; PMID:7793970
A;Accession: S66261
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-711 <KAB>
A;Cross-references: UNIPROT:Q47900; UNIPARC:UPI00000B2115; EMBL:D42121; NID:g577283; PID
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidyl-peptidase hydrolase

Query Match 20.4%; Score 791.5; DB 2; Length 711;
Best Local Similarity 30.3%; Pred. No. 8.6e-47;
Matches 216; Conservative 125; Mismatches 274; Indels 97; Gaps 24;

Qy 31 AEYGNSSVFLNSTFDFEGHSHINDYSISPDGQFILLIYVYVQWRHSYASYDIYDLNKR 90
Db 60 AKYSYTKSQEKKNIVD---GSFGYTFNSDESKILLQKSSQSIYRHSFLGKPEVKDLKSR 116

Qy 91 QLITBERIPNNTQW---TWSPVGHKLAYVWNNDIYVIEPNLPYRITWTGKEDIYNG 147
Db 117 TVVS-----LNNANWIEPKFSPDGSVAFIADNNLNFYQDLNTGKIQTITTDGKNEI 172
Qy 148 ITDWVVEEVEFSAYSALWSPNGTFLAYQFNDTEVPLIBYSYDESLOYKTVRPYP 207
Db 173 LGDWVYVEE--FGHADYVQWKNAGDALVFRFDRKVPINPIIYQW--LYPKLMTYKYP 229
Qy 208 KAGAVNPTKFFVWNTDSLSSTVNATSIQITAPASMLIGHYLCDDVTWAT--OERISLOW 265
Db 230 KAGEENSANV-----TAYLYQLSSGSAQLNFSS-----EKYIPQLFQTNANDEIVVAT 279
Qy 266 LRRIONYSVMDICD-----YDESSGRWNCILVARQHIEIMSTTGMWGRFRPSPHF 314
Db 280 ANRHQN--KVDLLKVTAKTAAVSKLFTETDNAM---IETDNLTWE----- 319
Qy 315 TLDGNSFYKIIISNEEGYRHICYFQIDKDCOTFITKGTWEVIGIEALTSYLYISNEYK 374
Db 320 FLDDNSFL-WASERDGRHLHYDYAAGKLKQKQVSKGDWEII-----NYGY 364
Qy 375 MPGGRNLYKIQLSDYTKVTCLSCEL--NPERCOY-----SVSFSKAEKYVQLRCSGP 425
Db 365 NPATKEVY-IQTTEKGSINKVSKLINTGKTQLLSNAEGNNSAASKTNYFINTSSTA 423
Qy 426 GLPLYTLHSSVNDKGLRVLEDNSALDKMLQVQVPSKKLDFIIL-----NETKFWYQMI 480
Db 424 KVPTKYLKANGKDVKEIQNDDLLKLSDNFIK--EFITIPNAAGQMNAM--MIK 479
Qy 481 PPHFDKSKYPLLLDVYAGCSQKADTVFRLN---WATYLASTENIIVASPDGRSGYQ 537
Db 480 PKNFDPACKYPVPMFYSGSQVANSWDGGNGIWFDMLAQ--KGYLVVCDVGRGTGFRG 538
Qy 538 DKIMHAINRRLGTFFVEVDQIEAARQFSKMGVDNKRITAIWGSYGGYVTSWVLGSGGV 597
Db 539 TTKYKVTYKNLGKYETEDQITAKWLGNSQSYVDKSRIGIFGYSYGGYMASLAMTKGADV 598
Qy 598 KCGIAPVSRWEYDVSVYTERYMGLTPTEDNLDHYRNSVTMSRAENFKQVEYLLIHGTA 657
Db 599 KMGIAVAPVTVNRFYDSIYTERF--LOTQENKDGVDLNSPTTYAKLLKG--KELLIHGTA 655
Qy 658 DNVHFPQSAQISKALVDVGDVFOAMWYTTDEDHGIIASSTAHOHIYTHMSHF 709
Db 656 DNVHFPQNSMERFSEALIQNKQDFDMAYDPDKNHSIIGGNTRPOLYKEMTNYI 707

RESULT 11
T19514
Hypothetical protein C27C12.7 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19514
R;Thomas, K.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19134
A;Accession: T19514
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-829 <WIL>
A;Cross-references: UNIPROT:Q18253; UNIPARC:UPI000007DCS3; EMBL:Z69883; PIDN:CAA93743.1
A;Introns: 4/3; 51/1; 123/3; 166/3; 188/2; 279/2; 392/3; 501/3; 553/1; 583/1; 606/2; 64
C;Superfamily: dipeptidyl-peptidase IV

Query Match 19.2%; Score 745; DB 2; Length 829;
Best Local Similarity 28.8%; Pred. No. 1.8e-43;
Matches 227; Conservative 128; Mismatches 282; Indels 150; Gaps 35;

Qy 8 YSLRWISDHEYLKQENNLVFNAEYGNSSVFLNSTFDFEGHSHINDYSISPDGQFILL 65
Db 105 YDWIPLPDGSFVQMDD----FTIRKQMKKIFLGSSVAEPFFN-----NGEVVKAL 151

```
QY 66 ---LEYNVVKQ-----WRHSYTSYDIYDLNKKQLITEERIPNNTQW-----V 105
Db 152 SSMVKYAGSKKNVLRWRSABEYLIVKINNTVST-----QHVGPENSLIQAF 204
QY 106 TWSP--VGHKLAVVNNNDIYVKLEPNLPSVRIWT--GKEDIYNGITDWYEEVFSAYS 162
Db 205 YWNPANSSNDFVYVHYNLYQKDPKPGAIQLTVGGSTFRFGLANWLYEEBIELEAS 264
QY 163 ALWMSRNGTFLAYAFQNDTEVPLIEYSFYSDLESQPKTVRVYPKAGAVNPT--VKFFV 221
Db 265 AVWSPSGRVYSLRFPDREVNRIFLPKYTD--SYVEPELVPKAGVQNNLTQYIW 323
QY 222 NTDLSLSVYVWATSIQITAPASMLIGDHYLCDVTWAT-----QERISLOWLRIONY 272
Db 324 DSENHKIVETAPPNELSAAN-----GDYYLTNKNWITMPRNGSDLGBERLVTWANDQNH 379
QY 273 SVMIDICDYDESSGRWNCILVA-----ROHIEWSTTGMVGRFRPSE-----PHTILD 317
Db 380 VYFSLCNEQD-----CVMALSFQFSDNRQLWVSPKDVGRGP--PTETGFLTVPKHKDD 432
QY 318 GNSFYKIISNEGYRHCYFQIDKDCFTFITKTWE-----VIG-----IEALT----- 361
Db 433 GNI-----YNHVAHVELDGTGTGKITWIGENFDVILVGYSSKIDALTPESAYGD 482
QY 362 --SDLYIYISNEYKMGPGGRNLYKQLSDYTKTCLSCELNPERCQY--SVSFSKEAYY 418
Db 483 GVGEFTYIVRE-----AMYSNKKTTLLQKVTD-----QPEDCKTLGQSADPTGQRI 529
QY 419 QLRCSGP--GLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVMPSSKKLDFIIL--NETKFW 475
Db 530 VVOCEKPFNDTRLVYV--DVVDTTKIMLEGGT--KAVIPFDVFNKMGKGLKLPDSIGDH 585
QY 476 YQMLPFPHPDKSKYKPLLLDVYAGPCS-----QKADTVFRLNWTATYLASTENIIVASDGR 531
Db 586 YWMLTPANLLDGAKIPLLLDIYGPDSKQVFKTPTAH---AIQIVSQYDIAYARIDVR 641
QY 532 GSGYQGDKIMHAINRLRGTFEVEDQIEAARQF--SKMGFVNDKRIATWGSYGVTSMVL 590
Db 642 GTGGRGMDVKEAVYRKLGDAAEVVDTLDMIRAFINTFTGFIDEDRIAVMGWSYGGFLTSKIA 701
QY 591 GSGSG--VFKGCIAPVPSRWEYVDSVTERYMGCLPTPENLDHRYNSTWMSRAENPKQVE 649
Db 702 IKDQGLVKCAISIAPTDFKYYDSAYTERYLG--QPAENLQYIINTVIPHARNTNVK 759
QY 650 YLLIHGTADDNVHFQOASQISKALVDVGVDFQAMWYTDDEHGFIASSTAHO--HIYTHMSH 707
Db 760 YLLAHGERDDNVHYQNSARWSEALQNGIHFTQLVYANEAH-----SLSHKGLFHYLGEVQR 815
QY 708 FI--KQCF 713
Db 816 FLNDCGF 822

RESULT 12
dipeptidyl aminopeptidase (EC 3.4.14.-) - Yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O5045; protein YOR219c; protein YOR50-9
C;Species: Saccharomyces cerevisiae
C;Date: 03-Mar-1994 #sequence revision 03-Mar-1994 #text_change 09-Jul-2004
C;Accession: A49737; S45451; S60946; S67112; S71721
R;Santa Anna-A, S.; Herskowitz, I.
A;Reference number: A49737
submitted to the Protein Sequence Database, July 1993
A;Molecule type: DNA
A;Residues: 1-931 <SAN>
A;Cross-references: UNIPROT:P33894; UNIPARC:UPI0000136060; GB:L21944; NID:g347196; PIDN:
R;Anna-Arriola, S.S.; Herskowitz, I.
Yeast 10, 801-810, 1994
A;Title: Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase
A;Reference number: S45451; MUID:95066382; PMID:7975897
A;Accession: S45451
```

```
A;Molecule type: DNA
A;Residues: 1-931 <ANN>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:g347196; PIDN:AAA35119.1;
R;Gallison, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A;Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome
A;Reference number: S60938
A;Accession: S60946
A;Molecule type: DNA
A;Residues: 1-931 <GAL>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1;
R;Boyer, J.; Fairhead, C.; Gaillon, L.; Gallison, F.; Michaux, G.; Thierly, A.; Dujon, B.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67104
A;Accession: S67112
A;Molecule type: DNA
A;Residues: 1-931 <BOY>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:Z75127; NID:g1420507; PIDN:CAA99437.1;
R;Gallison, F.; Dujon, B.
Experimental source: strain S288C
Yeast 12, 877-885, 1996
A;Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV
A;Reference number: S71713; MUID:96437977; PMID:8840505
A;Accession: S71721
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-931 <GAW>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1;
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGD:STE13; YC11
A;Cross-references: SGD:S0005745; MIPS:YOR219c
A;Map position: 15R
C;Function:
A;Description: involved in processing of alpha-factor prepropheromone
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
F;113-150/Domain: transmembrane #status predicted <TM>
F;377/Binding site: carbohydrate (Aen) (covalent) #status predicted
F;785,863,896/Active site: Ser, Asp, His #status predicted

Query Match 18.7%; Score 726; DB 2; Length 931;
Best Local Similarity 26.9%; Pred. No. 4.4e-42;
Matches 195; Conservative 146; Mismatches 318; Indels 66; Gaps 23;

QY 29 FNAEYGNSSVFLSENSTDFEFGHSINDYSISDPGQFILLEYNVYKQMRHYSYASDIYDLN 88
Db 223 FEVNLGGRFLYEGVEFTVSTVQIN--YKLDK----LIFGTNLESEPRHSSKGFYWKLDN 277
QY 89 K--RQLITEERIPNNTQ-----WVTWSPVGHKLAVYWNNDIYVYKIEPNLPSYRITWT 138
Db 278 TGNIEPILPPEKSDDNDELGLSKLSYAHFSPAYNTIYFVYNNLFLQQVNSGVAKKVTED 337
QY 139 GKEDIYNGITDWYEEVFSAYSALWNSPNCGTFLAYAQFNDTEVPLIEYSFYSDLESQY 198
Db 338 GSKD--IFNAKPDWIEVEEVLASDQAIWAPDDSKAVAFRNDTSVDDIRLNRYTNWNEAY 396
QY 199 PXTVRVPYKAGAVNPTVKFFVVNTDSLSSVTNATSIQITAPASMLIGDHYLCDVTWATQ 258
Db 397 LSDTKIKYKPKGQFQNPQDFLNV-----LQNGIYSINTGGQK---DSILYNGKWISP 447
QY 259 ERISLOWLRIONYQVMDICDYDESSGRWNCILVARQHITEMSTGWGRFR-----PSEPH 313
Db 448 DTFRFEITDR--NSKILDVKVYDIPSSQ--MLTVRNTNSLNFNGWIEKTKDILSLPPKPE 503
QY 314 FTLDGNSFYKIISNEEGYRHCYF--QIDKDCFTFITKTWEV-----IGIEALTSDLYY 367
Db 504 LKRMVGYDIDHADSRGSHLFYFTVPAKEPIQLTKGNWETGNGIVGEYET--DTIFF 562
QY 368 ISNEYKMGPGGRNLYKQLSDYTKTCLSCELNP--ERCQYVSVSFSKEAKYVQLRCSGP 426
Db 563 TANEIGVM--SQHLYSISLSTDTTQNTFQSLQNPDKYDFDFELSSSARYAISKKLGP 620
QY 427 LPYLTILH-----SSVNDKGLRVLEDNSALDKMLQNVMPSSKKLDFIILNE--TKFWYQ 477
```

Db 621 TPIKAVAGPLTRVLNVAEIHDDSIQLTKDEKFKIKNYDLPTSYKTMVLDDGVEINYI 680
QY 478 MILPPHFDKSKYPLLDVTVAGPCSKADTVFRLNWTATYLASTENIIVASFDFGRSGYQG 537
Db 681 EIKPANLNPKKKYPILNIIYGGPGSQFTTKSSLAFFQAVVSGLDVIVLQIEPRGTGGK 740
QY 538 DKIMHAINRRLGTFEVEDQIEAARQFSGKWG--FVDNKRIAIWGWSYGYVTSMVLGSGG 595
Db 741 WSPRSWAREKLYGWEPRDITVTKKFIQRNSQIHDESKIAIWGWSYGGFTSLKTVELDNG 800
QY 596 -VFKCGIAPVPSRWYDYVYTERYMGFLPTPEDNLDHYRNSVMSRAENFKQVEYL--- 651
Db 801 DTFKYAMAVAPVTNTWLYDSYTERYMNQPS-ENHEGYFEVSTI---QNFKSPESLKRL 855
QY 652 -LHGTADDNVHFOQSAQISKALVDGV-DFOAMWYTDDEHGIIASSTAHOHIYTHMSHFI 709
Db 856 FIVHGTDDNVHIOHTFRVDQLNLLGLTNYDMHIFPDSDHSIRYHNAQRIVFQKLYYL 915
QY 710 KQCFS 714
Db 916 RDAFA 920
RESULT 13
T25173
hypothetical protein T23F1.7a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25173
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19990
A;Accession: T25173
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-779 <WIL>
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN: CAB03411.1
A;Experimental source: clone T23F1
C;Genetics:
A;Gene: CRSP:T23F1.7a
A;Map position: 5
A;Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1
C;Superfamily: dipeptidyl-peptidase IV
Query Match 18.6%; Score 720; DB 2; Length 779;
Best Local Similarity 27.6%; Pred. No. 8.8e-42;
Matches 218; Conservative 124; Mismatches 296; Indels 152; Gaps 33;
QY 2 TYRLKLYSLRWISDHELYKQ---ENNILVFNA-----EYGNSSVFLNSTFDFEGH 50
Db 60 SWRLNVSDLRSL--RYPYAEPAFTDNNAVMQSWEGVEIVDGVSRLLI-----FGR 108
QY 51 SINDYSISPDGQ---FILLYN-----YVQWRHSYTA SYDIYDLNKRQLITE 95
Db 109 E-NGAETPSADRKYFAMMDHAPNPGWNPQNETFHLKIVNNNERITYDI-GLAKEESV-- 164
QY 96 ERIPNNTQWTVSPGVGHKLAYVWNNDIYVKIEBNLPSYRITWTGKEDIYNGITDWMYEE 155
Db 165 -----IQAFKWKNGKFNDFVVESENKIYQSSPEEGLTRVNSGGEHTV-DGLFDWYEE 217
QY 156 EVFSAYSALWSPNGTFLAYAQFNDTEVPLIEYSFYS-----DESLOYPKTVRPYYPK 208
Db 218 EIFGRKDAWMWSTKGDQLAYASDNHLTKNVSLKTYHRLPEYPIDTNFHPKPT----- 271
QY 209 AGAVNPVKPFVYNTDLSLSTVNTATSIQITAPASMLIGDHYLQDVTW-----ATQERISLQ 264
Db 272 -AKVLPY-----TLSIWNKKTQSRQLDVQLKDSLSHYLLAVKWLINGTEQLVSV- 323
QY 265 WLRRIQNSYWDICDYDESSGRNCLVARQHIE--MSTTGWGRFRPSPHFTLDGNSFY 322
Db 324 WTRNYQNEVALTICD-----WDTAICRLEFEYKYKASKRWVTH-----DDFH 364

QY 323 KIISNBEGYRHICVFOI--DKKQCT-----PITGKTWEVIGIEALTS 362
Db 365 SITSPED-----TUFFLLPHDRNNAFOQVAVSLRSLSHGQLRTPKFLNLTGEVDVTSINGINK 420
QY 363 DYLYIYSNEYKMGPGGRNLKYIQLSDYTK--VTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 421 ETRTIFPHAAAPKPSHRSLFSYSLADESRNSAYCISCSI--KNCTWAAQAMDDQMKTAIV 478
QY 421 RCGSGPLPLYLTHSSV-----NDKGLRVLEDNALSADKMLQNVQMPKSLFILLNEKPF- 474
Db 479 SCKGPAAP-----HTAIVNLTRMDSDKKTEHANLLYDKTYQN-RVEEAGLPVIIKETIKIS 533
QY 475 -----WYQMLPPIHF--DKSKKYPILLDYYAGPCSKADTVFRLNWTATYLASTENIIVA 526
Db 534 DDFDALIKLSIPKDIYNRDGHQAIPLIHVYVYGFNDQNTKEATQIGITEEVVASASQAAIL 593
QY 527 SFGRSGVGGDKIMHAINRRLGTFEVEDQIEAARQFSGKM--GFVDNKRIAIWGWSYGGY 584
Db 594 RIDGRSGGRGWKYRSALYQGLTVEVEDQIKAVVLRVLRHLLDARRVAVFGWSYGGF 653
QY 585 VT-SMVLGSGGVFKCGIAPVPSRWYDYVYTERYMGFLPTPEDNLDHYRNSVMSRAE 643
Db 654 MTLSMYNEAPEQFFKCAVSPVTFNFAYYDATYTERYMG---DAPLESY--SDVTKKLD 707
QY 644 NFKQVEYLLIHGTADDNVHFOQSAQISKALVDGVDFQAMWYTDDEHGIIASSTAHOHIYT 703
Db 708 NFKSTRLLMHGILLDDNVHFPNSAILDELQNRGVDFDLVMYPNQAHLSRSTS--HWVG 765
QY 704 HMSHFIKQCF 713
Db 766 KMTFLRQCF 775
RESULT 14
T25174
hypothetical protein T23F1.7b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25174
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19990
A;Accession: T25174
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-799 <WIL>
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN: CAB03412.1
A;Experimental source: clone T23F1
C;Genetics:
A;Gene: CRSP:T23F1.7b
A;Map position: 5
A;Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1
C;Superfamily: dipeptidyl-peptidase IV
Query Match 18.5%; Score 716; DB 2; Length 799;
Best Local Similarity 27.5%; Pred. No. 1.7e-41;
Matches 221; Conservative 122; Mismatches 300; Indels 162; Gaps 33;
QY 2 TYRLKLYSLRWISDHELYKQ---ENNI-----LVFNABSYG----- 34
Db 60 SWRLNVSDLRSL--RYPYAEPAFTDNNAVMQSWEGVEIVDGVSRLLIFGRENGAETP 116
QY 35 -----NSSVFLNSTFDFEGHSINDYSISPDGQFILLIYNYVVKWRHSYTA 80
Db 117 SADRKYFAMMDHAPNPGWNPQNETF-HLKIVNNNERLNP-----LLPFE-VEELFRELSD 169
QY 81 SYDIYDLNKRQLITEERIPNNTQWTVSPGVGHKLAYVWNNDIYVKIEBNLPSYRITWTGK 140
Db 170 SRITYDIGLRK---EESV---IQAFKWKNGKFNDFVVESENKIYQSSPEEGLTRVNSGG 223
QY 141 EDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFNDTEVPLIEYSFYS-----D 193
Db 224 EHTV-DGLFDWYEEEIFGRKDAWMWSTKGDQLAYASDNHLTKNVSLKTYHRLPEYPID 282

Qy 194 ESLOYPKTVRPYPKAGAVNPTVKFVVVNTDSJSSVTNATSIQTTPASMLIGDHYLCDV 253
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 283 TNFHYPKTF-----AKVLPTY-----TLISWNKKTEQSRLQDLVDQLKDSLVSYHLLAV 329
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 254 TW-----ATOERISLOWLRRITONYSVMIDICDYDESSGRWNCIVARQHIE--MSTTGWGRF 307
| :
Db 330 KWLSEINGTEQLVSV-WTNRYONEVALTICD-----WDTAICRLEFEFYKYASKRWVTH- 380
| :
Qy 308 RPSPEPHTLDGNSFYKIISNEEYRHICYFOI--DKKDCT-----FI 347
| :
Db 381 -----DDPHSITSFD-----TLFELLPHDXEDNAFQQVASLRLSHGQLRTPKFL 425
| :
Qy 348 TKGTFWEVIGIEALTSDFLYYISNEYKMGPGRMNLKYTLQSDYTK--VTCLSGELNPERCQ 405
| :
Db 426 NLGEYDVTSINGINKETRTIFFHAAAPKPSHRSLSFSYSLADESRNSAYCISCISI--KNCT 483
| :
Qy 406 YYSVSFSKEAKYOLRCSGPLPLYTLHSSV----NDKGLRVLEDNSALDKMLQNVOVP 460
| :
Db 484 WAQAOMDDQMKTALVSCCKGAAP-----HTAIVNLTMRMDSKKKTEHANLLYDKTYQN-RVE 538
| :
Qy 461 SKKLDFILNETKF-----WYOMILPPHF-----DKSKYPPLLDDVYVAGPCSQADTVFRL 511
| :
Db 539 EAGLPVIKETIKISDDFDALIKLSIPKOIYNRDKHQAIPLIHVHYGGPNQNTKEATQI 598
| :
Qy 512 NWATYLASTENIIIVASDFGRSGVGQDGKIMHAINRLTGTFEVEDQIEAARQFSKM--GFV 569
| :
Db 599 GIEEVVASAQAAITLRIDGRSGRGWKYSALYQGLGTVEVEDQIKAIKVVLRLYRHLL 658
| :
Qy 570 DNKRIAIWGSYGGYVT-SMWLGSGGVFKGIATAVPVRSWEYYVDSVYTTERYMGLPTPED 628
| :
Db 659 DARVAVFWSYGGFWLMSWNEAPEOFFKCAVSAVPTNFAYYDATYTERYNG-----DA 714
| :
Qy 629 NLDHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFQOOSAQISKALVDVGVDFOAMWYTD 688
| :
Db 715 PLESY--SDVTXKLDNFKSTRLLLMHGLLDDNVHFQNSAILIDELQNRGVDFDLWVYPNQ 772
| :
Qy 689 DHGIASSTAQHIYTHMSHIKQCF 713
| :
Db 773 AHSLSRSTS--HVVGRMTHTFLROCF 795
| :

RESULT 15

A87516
dipeptidyl peptidase IV [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: A87516
R/Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
n., J.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; UID:21173698; PMID:11259647
A/Accession: A87516
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-738 <STO>
A/Cross-references: UNIPROT:Q9A6S0; UNIPARC:UPI00000C7616; GB:AEO05673; NID:g13423647; H
C/Genetics:
A/Gene: CC2154

Qy	169	NGTFLAYAQFNDTEVPLEIYSFYDESIOYPKTVRVVPYKAGAVNPVTYKFFVVNTDSLSS	228
Db	216	DESRIYVTRVDESVDIVP---RADIGPGGATVWQRYPRAGRPNAVVDLF-----	263
Qy	229	VTNATSIQITAPASKMLIGDHYLCDVTW--ATQERISLOWLRIONYSVMIDICDYDESSGRW	287
Db	264	VRDLASGKVATLDLGANKDIYVARVAMSADGKTIVVQRLSRDQ--KTJLDLAFDAATGAG	321
Qy	288	NCLVARQHIEMSTTGWGRPRPSPBPHFTLDGNSFYKII-----SNEEYGRHICYFOID	340
Db	322	KTILT-----DTPHFTEVSNDRFPLTDTGTFLWGSBKDGQHLYRYAAD	365
Qy	341	KKDCFTFKGTWEGVIGIEALTSYLYIISNEYKMGPGGRNDLYKIQLSYTKVTCLSCELN	400
Db	366	GKLIATIKTGDWPVIGLEG-----VDEARKVAIPESASID	399
Qy	401	-PERCOYYSVSFSEANKYYQLRCSG-----PGLPLYTLHSSVNDK	439
Db	400	TPIERRLYEVSAYKPGKPKALTSAGGWWAAKVADNGGAFAGTYSDPKTPSQOTALYSADGK	459
Qy	440	GLRVLEDNSALD-----KMLQNVQWPS-----KKLDFIILNETKFWYQWILPFPKSKKY	490
Db	460	RVRWIEENKLAEGHPYPPYAANLPQPEPGSLKAAD-----GET-LHYEILKPIGFDPACKY	514
Qy	491	PLLLDVYAGPCSQKADTVFRLNW-----ATYLASTENIIVASFDRGSGYQGDKIMHAIN	545
Db	515	PAIVSVYGGPHAQRV---MKNWHSPSBRTYLEA--GYVIFKLDNRGSGNRSAKFMRALD	568
Qy	546	RLRGTFEVEDQIEARQFSKMGFVDNKBRIAIGHWSYSGGVVTSWVLGSSGGVFKCIGIAPAP	605
Db	569	RKLGTFEVEDQLLGAKFASQPYVDADKLGVMGWSYGGFMALMLLTAEINTFPKGAAGAAP	628
Qy	606	VSRWEYDVSVTRYMGLTPEDNLDHNRNSTVMSRAENPQKQVEYLLJHGTADDNVHFQQ	665
Db	629	PTESLVDYATERYMG--KDENKAGAYSDINNRIDKLPAGSLLLLHGHVADDNVIFEN	686
Qy	666	SAQISKALVDVGVDFQAMWYTDEDHGIASSTAHH--QHITYTHMSHF	708
Db	687	STRJMAALORKAIIIPENAWYPGERSHSPGSKTKGLSVLTKTHLDPF	731

Search completed: January 30, 2006, 14:23:27
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 23:58:33 ; Search time 11 Seconds
(without alignments)
704.874 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877

Sequence: 1 NTYRLKLYSLRWISDHELYL.....AQHIYTHMSHFYKQCSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA New:*
- 1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
 - 2: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
 - 4: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
 - 5: /cgn2_6/ptodata/2/pubppaa/US05_NEW_PUB.pep:*
 - 6: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
 - 7: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep:*
 - 8: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	766	6	US-10-522-789-2
2	3834	98.9	762	7	US-11-116-939-13
3	2138	55.1	760	7	US-11-186-284-55
4	525	13.5	882	7	US-11-151-601-20
5	272	7.0	657	7	US-11-179-977-1
6	124	3.2	668	6	US-10-454-437-118
7	115	3.0	1593	6	US-10-453-372-1092
8	112.5	2.9	1436	6	US-10-453-372-1094
9	110	2.8	19	7	US-11-116-939-27
10	110	2.8	72	7	US-11-151-601-22
11	109.5	2.8	877	7	US-11-077-550-157
12	109	2.8	2516	6	US-10-647-956A-2
13	107	2.8	866	7	US-11-077-550-32
14	107	2.8	871	7	US-11-077-550-155
15	107	2.8	873	7	US-11-077-550-163
16	107	2.8	873	7	US-11-077-550-165
17	107	2.8	876	7	US-11-077-550-66
18	107	2.8	878	7	US-11-077-550-62
19	107	2.8	879	7	US-11-077-550-30
20	107	2.8	908	7	US-11-077-550-64
21	107	2.8	914	7	US-11-077-550-60
22	107	2.8	949	7	US-11-077-550-68
23	107	2.8	1130	7	US-11-077-550-44
24	107	2.8	1132	7	US-11-077-550-46
25	106	2.7	878	7	US-11-077-550-12

26	106	2.7	907	7	US-11-077-550-16	Sequence 16, Appl
27	106	2.7	953	7	US-11-077-550-14	Sequence 14, Appl
28	106	2.7	1013	7	US-11-077-550-18	Sequence 18, Appl
29	106	2.7	4495	6	US-10-453-372-1002	Sequence 1002, Ap
30	105.5	2.7	871	7	US-11-077-550-2	Sequence 2, Appli
31	105.5	2.7	871	7	US-11-077-550-8	Sequence 8, Appli
32	105.5	2.7	871	7	US-11-077-550-26	Sequence 26, Appl
33	105.5	2.7	871	7	US-11-077-550-153	Sequence 153, App
34	105.5	2.7	873	7	US-11-077-550-6	Sequence 6, Appli
35	105.5	2.7	873	7	US-11-077-550-149	Sequence 149, App
36	105.5	2.7	873	7	US-11-077-550-151	Sequence 151, App
37	105.5	2.7	873	7	US-11-077-550-167	Sequence 167, App
38	105.5	2.7	873	7	US-11-077-550-169	Sequence 169, App
39	105.5	2.7	879	7	US-11-077-550-159	Sequence 159, App
40	105.5	2.7	894	7	US-11-077-550-4	Sequence 4, Appli
41	105.5	2.7	1127	7	US-11-077-550-40	Sequence 40, Appl
42	105.5	2.7	1129	7	US-11-077-550-42	Sequence 42, Appl
43	105.5	2.7	3003	6	US-10-453-372-1080	Sequence 1080, Ap
44	105.5	2.7	3361	6	US-10-453-372-1082	Sequence 1082, Ap
45	104.5	2.7	658	6	US-10-873-528-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1

US-10-522-789-2
; Sequence 2, Application US/10522789
; Publication NO. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-522-789-2

Query Match	100.0%	Score 3877	DB 6	Length 766
Best Local Similarity	100.0%	Pred. No. 2.3e-299		
Matches 716	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	NTYRLKLYSLRWISDHELYLKOENNLVFNNAEYGNSSVFLNENSTFDFGHSINDYSISPD	60	
Db	51	NTYRLKLYSLRWISDHELYLKOENNLVFNNAEYGNSSVFLNENSTFDFGHSINDYSISPD	110	
Qy	61	GQPILEYNVKQWRHSYASDYIDLNKRQLITEERIPNNTQVMTWSPVGHKLAYVWNN	120	
Db	111	GQPILEYNVKQWRHSYASDYIDLNKRQLITEERIPNNTQVMTWSPVGHKLAYVWNN	170	
Qy	121	DIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVFSAYSAIHWSPNCTFLAYAQFND	180	
Db	171	DIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVFSAYSAIHWSPNCTFLAYAQFND	230	
Qy	181	TEVPLLEYFSDSESLQYPKTVRVPYKAGAVNPTVKFFVWNTDSLSSVTNATSIQITAP	240	
Db	231	TEVPLLEYFSDSESLQYPKTVRVPYKAGAVNPTVKFFVWNTDSLSSVTNATSIQITAP	290	
Qy	241	ASMLIGHDYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRNCLVARQHIEHMT	300	
Db	291	ASMLIGHDYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRNCLVARQHIEHMT	350	
Qy	301	TGHWGFRPSEPHFTLDGNSFYKLIISNEGYRHI CYFQDKDKDCTITTKGTWEVIGIEAL	360	
Db	351	TGHWGFRPSEPHFTLDGNSFYKLIISNEGYRHI CYFQDKDKDCTITTKGTWEVIGIEAL	410	

QY 361 TSDYLYYSNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAYQOL 420
 DB 411 TSDYLYYSNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAYQOL 470
 QY 421 RCGSPGLPLYLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFWQMIL 480
 DB 471 RCGSPGLPLYLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFWQMIL 530
 QY 481 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTYLASTENIIVASFDGRGSGYQGDKI 540
 DB 531 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTYLASTENIIVASFDGRGSGYQGDKI 590
 QY 541 MHAINRLRGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYTSVWLGSGGVFKCG 600
 DB 591 MHAINRLRGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYTSVWLGSGGVFKCG 650
 QY 601 IAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
 DB 651 IAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
 QY 661 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIFKQCFSLP 716
 DB 711 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIFKQCFSLP 766

RESULT 2
 US-11-116-939-13
 ; Sequence 13, Application US/11116939
 ; Publication No. US20050265995A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stephen Tomlinson
 ; APPLICANT: Richard J. Quigg
 ; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
 ; FILE REFERENCE: 19113.011502
 ; CURRENT APPLICATION NUMBER: US/11/116,939
 ; PRIOR FILING DATE: 2005-04-28
 ; PRIOR APPLICATION NUMBER: 60/565,907
 ; PRIOR FILING DATE: 2004-04-28
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 762
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
 ; OTHER INFORMATION: construct
 US-11-116-939-13

Query Match 98.9%; Score 3834; DB 7; Length 762;
 Best Local Similarity 98.6%; Pred. No. 5.8e-296;
 Matches 711; Conservative 0; Mismatches 6; Indels 6; Gaps 1;
 QY 2 TYRLKLY-----SLRWISDHELYKQNNILVFNAEYGNSSVFNLENSTDFEFGHSINDY 55
 DB 42 TYTILTYLKNYKLLRWISDHELYKQNNILVFNAEYGNSSVFNLENSTDFEFGHSINDY 101
 QY 56 SISPDGQFILLEYYNKKQMRHSYASYDIYDLNKRQLITEERI PNNTQWYTWSPVGHKLA 115
 DB 102 SISPDGQFILLEYYNKKQMRHSYASYDIYDLNKRQLITEERI PNNTQWYTWSPVGHKLA 161
 QY 116 YVWNNDIYVKIBNLPSPYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAY 175
 DB 162 YVWNNDIYVKIBNLPSPYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAY 221
 QY 176 AQFNDTEVPLIEYSFYSDESQYKPTVRVPYKAGAVNPTVKFVNTDSLSSVTNATSI 235
 DB 222 AQFNDTEVPLIEYSFYSDESQYKPTVRVPYKAGAVNPTVKFVNTDSLSSVTNATSI 281
 QY 236 QITAPASMLIGDHYLCDDVTWATERISLOWLRRIQNYVMDICDYDESSGRMNCVLARQH 295
 DB 282 QITAPASMLIGDHYLCDDVTWATERISLOWLRRIQNYVMDICDYDESSGRMNCVLARQH 341

QY 296 IEWSTTGWGRPRPSEPHEFTLDGNSFYKLIISNEEGYRHCYFQIDKDKCCTFITKGTWEVI 355
 DB 342 IEWSTTGWGRPRPSEPHEFTLDGNSFYKLIISNEEGYRHCYFQIDKDKCCTFITKGTWEVI 401
 QY 356 GIEALTSDYLYYSNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEA 415
 DB 402 GIEALTSDYLYYSNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEA 461
 QY 416 KYQJLRCSGPGPLYLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFW 475
 DB 462 KYQJLRCSGPGPLYLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFW 521
 QY 476 YQWILPPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTYLASTENIIVASFDGRGSGY 535
 DB 522 YQWILPPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTYLASTENIIVASFDGRGSGY 581
 QY 536 QGDKIMHAINRLRGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYTSVWLGSGSG 595
 DB 582 QGDKIMHAINRLRGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYTSVWLGSGSG 641
 QY 596 VFKCGIAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHG 655
 DB 642 VFKCGIAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHG 701
 QY 656 TADDNVHFOQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIFKQCFSL 715
 DB 702 TADDNVHFOQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIFKQCFSL 761
 QY 716 P 716
 DB 762 P 762

RESULT 3
 US-11-186-284-55
 ; Sequence 55, Application US/11186284
 ; Publication No. US20050266493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: Berger, Allison
 ; APPLICANT: Guillemette, Tracy L.
 ; APPLICANT: Kamatkar, Shubhangi
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Thibodeau, Stephen N.
 ; APPLICANT: Burgart, Lawrence J.
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
 ; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; TITLE OF INVENTION: THERAPY OF COLON CANCER
 ; FILE REFERENCE: MPW01-029P2RNM
 ; CURRENT APPLICATION NUMBER: US/11/186,284
 ; CURRENT FILING DATE: 2005-07-21
 ; PRIOR APPLICATION NUMBER: US/10/301,822
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 60/339,971
 ; PRIOR FILING DATE: 2001-12-10
 ; PRIOR APPLICATION NUMBER: US 60/361,978
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 60/381,988
 ; PRIOR FILING DATE: 2002-05-20
 ; NUMBER OF SEQ ID NOS: 228
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 55
 ; LENGTH: 760
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-11-186-284-55

Query Match 55.1%; Score 2138; DB 7; Length 760;
 Best Local Similarity 53.3%; Pred. No. 1.3e-161;
 Matches 383; Conservative 134; Mismatches 187; Indels 14; Gaps 7;
 QY 2 TYRLKLYSLRWISDHELYKQ-ENNILVFNAEYGNSSVFNLENSTDFEFGHSIN--DYSIS 58

Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-11-551-601-22

Query Match 2.8%; Score 110; DB 7; Length 72;
Best Local Similarity 48.3%; Pred. No. 0.022;
Matches 29; Conservative 3; Mismatches 28; Indels 0; Gaps 0;

QY 524 IVASFDGRGSGYQGDKIMHAINRRLGTFPEVEDQIEAARQFSKMGFVNDKRIAIWGWSYGG 583
DB 11 IYAVVDIRGGEGYQKQWHEAGTRKLKKNFDFIAAAEYLSKLYTSPKRIAIFGSNGG 70

RESULT 11

US-11-077-550-157
; Sequence 157, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 157
; LENGTH: 877
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-157

Query Match 2.8%; Score 109.5; DB 7; Length 877;
Best Local Similarity 19.5%; Pred. No. 0.73;
Matches 144; Conservative 91; Mismatches 259; Indels 243; Gaps 32;

QY 30 NAEYGNSSV---FLENSTDFEGHISINDYSISPDGQFILLEYNVVKWRHSYASVDIYD 86
DB 250 NAYEYSGLEVSFEELRTFG--GHDKAFIDSLQENEFRLYYNKKF-----DIAS 297

QY 87 -LNKRQLITEERIPNNQTQWTSVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIY 145
DB 298 TLNKAQSI-----GTTASIQYKQVFKPKYLLSEDTSGKFSVDKLF-----DKLY 344

QY 146 NGITDWYEEVSAYSALWNSNGTFLAY--RQFNDTEVPLIEYSPFSDIESLOYKTVR 203
DB 345 KMLTE-IYTEDNFVKFKV--LNRKTYLNPDKAVFKINIVPKVNYTIYDGFNLRLNML-- 399

QY 204 VPYPKAGAVNPTVKFFVNTDLSSTVNTATSIQITAPASMLIGHY--LCDVTWATQERI 261
DB 400 -----AANFNGQNTENNWNFTKLNFTG-----LFEFYKLLCVRGIIITSYTK 442
QY 262 SL--QWLRRIONYSVMDICDYDESSGRWNCCLVARQHIEMSTTGWVGRPRPSPHPHTLQNS 320
DB 443 SLDKGYNKIEGRALNDLC-----IKVNNMDLFFSPSEDNFTNDLNK 483
QY 321 FYKIISN---EEGVRHICYFQIDKDOCTEITKTGWEVIGIEALTSYDL--YYISNEVKG 375
DB 484 GEEITSDTNIETAAEENISLDLIQOYILTFNFDNEPENISIEENSSDIIGOLELMPNTERF 543
QY 376 PGRNLYKIQLSDYTKVTCISCELNPERCOYYSVSFSKEAKYQLRCGSGPLPLYTLHSS 435
DB 544 PNGK---KYELDXYTWPHYL-----RAQEPHEGKSRIA-----LTNS 577
QY 436 VNDKGLRVLEDNSALDKMLQNVQMPKSLDPIILNETKFWYQMIPLPHFKSKYPLLLD 495
DB 578 VNEALLN-----PSRVYTFP----- 592
QY 496 VYAGPCSQKADTVPRLNWAT-----YLASTENIIVASPDGRSGVQGDKIMH----- 542
DB 593 -----SSDYVKKVKNKATEAAMFLGWVEQLVDFDTESEVSTTDKIADITIIIPYIG 644
QY 543 -AINRRLGTPEVEDQIEAARQFS-----KMGFVNDKRIAIWGWSYGGYVTSWVLGSGSV 596
DB 645 PALN--IGNMLYKDDFVGALIFSGAVILLEFIEPIALPVLGTALVSYIANKVLTVOT-- 700
QY 597 FKCGIAVAPVSRWYVDSVYTERYMGLPTEPDNLHYRNSVMSRAENFKQVEYLLHGT 656
DB 701 ----IDNALSKRNEKWDVY-----KYIVTNWL 724
QY 657 ADDNVH-----FQSAQISKALVDVGVDFQAMWYTDDEHG-----IASSTAH 698
DB 725 AKYNTQIDLRKQKKEALENAEATKAI-----INYQNTYTEBEKKNINFNIDDLSSKLN 780
QY 699 QHI---YTHMSHFHKQC 712
DB 781 ESINKAMININKELNQC 797

RESULT 12

US-10-647-956A-2
; Sequence 2, Application US/10647956A
; Publication No. US2005025187A1
; GENERAL INFORMATION:
; APPLICANT: French-Constant, Richard
; APPLICANT: Bowen, David
; APPLICANT: Rocheleau, Thomas
; APPLICANT: Waterfield, Nicholas
; TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS
; FILE REFERENCE: 61645
; CURRENT APPLICATION NUMBER: US/10/647,956A
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: CURRENT APPLICATION NUMBER: US/09/817,514
; PRIOR FILING DATE: CURRENT FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: US 60/191806
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 2516
; TYPE: PRT
; ORGANISM: Photorhabdus luminescens
US-10-647-956A-2

Query Match 2.8%; Score 109; DB 6; Length 2516;
Best Local Similarity 18.5%; Pred. No. 3.4;
Matches 137; Conservative 89; Mismatches 205; Indels 310; Gaps 39;

QY 20 YKQENNILV--FNAE-----YGNSS-----VPLENSTDFEGHISINDYSISPDGQFILLE 67
DB 1239 YQGEDTLLVMFYNQDQLDSYKNVSMQGLYIFADMASKDMTPQSNVYR---DNSYQQFD 1295


```

QY 68 YNVYKWRHSYASDI-YOLNKRQ-----LITEERIPNNTQWVTWSPVGHKLAY 116
Db 1296 TNNVRNNRNYAEDEIYESSVSRKQYGMGDDYLSMVGNDIPT-----INXKAA- 1345
QY 117 VWNNDIVVKEPNLPSYRITWTGKEDIYNGITDWWYE-----EEVFSA 160
Db 1346 ---SSDLKIYISPKL-----RIHNG-----YEGQKRNQCNLMKYKGLGDKPIV 1387
QY 161 YSALWSPNGT---FLAYAQFNDETEVPLIEYS-----FYSDESLOYPKTVR-- 203
Db 1388 YTSLVGNPNNSNKLMEY-----PVYQYSGNTSGLNQRLLFHRDITI--YPSKVEAW 1437
QY 204 VPPKAGAVNPT-----VKPFVNTDSSLSSVTNATS-IQIT---APAS 242
Db 1438 IPGAKRSLTNQNAAGDDYATDSLNLKPDLDLQYIFMTDSKGTATDVSGPVEINTAISPAK 1497
QY 243 MLI-----GDHYLCDVTWATERISLQ-----WLRRIQNVSMDI- 277
Db 1498 VQIIVKAGGK---EQFTADKDVSIQPSFDEMNQYFNALEIDGSGLNFINNASIDVT 1554
QY 278 -----CDYDESS-----GRMNCLVARQHI 296
Db 1555 FTAPAEGRKLGYESPSIPVTLKVSTDNALTLLHNENGAQYMQWQSVYRTRLNLFARQLV 1614
QY 297 EMSTTG-----WVGRFRPSPHFTLDGNSFY-----KIISN 327
Db 1615 ARATTTGDTILSMETQIOBPQL---GKGYATFVIPPYNLSTHGDERNFKLYIKHVDN 1671
QY 328 EGVYRHCY-----FQID-----KKDCTF---ITKGTW-----E 353
Db 1672 NS---HIYSGQLTDTNITLFIPLDDVPLNQDYHAKVYMTFKKSPSDGTWNGPHFVRD 1728
QY 354 VIGIEAL-----TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPER 403
Db 1729 DKGIVTNPXKSLTHPESVNVNLNISSEPMDFSGANSLYFWELFYITPMLVAQRLLEHQ- 1787
QY 404 QYYSVSFSKEAYQYLCRSGGLPLPYTLHSSVNDKGLRV---LEDNS----- 448
Db 1788 -----NFDENRNLKYVWSPSG---YIVHGQIQNYQWVRPLEDTSWNSDPLSDVDPD 1838
QY 449 ---ALDKMLQNVQMPKSLDFI1-----LNETKFWYQWILPPHFDKSKYK 491
Db 1839 AVAQHPDMHYKVTFMRTLDLLIARGDHAYRQLERDTLNEAKWYQWAL---HLLGDKPYL 1896
QY 492 LLLDVYAGPCSQKADTVFRNLN 512
Db 1897 PLSTWSDPRLDRAADITQN 1917

```

RESULT 13

```

US-11-077-550-32
; Sequence 32, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22

```

```

; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 32
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-32

```

Query Match 2.8%; Score 107; DB 7; Length 866;
 Best Local Similarity 19.6%; Pred. No. 1.1;
 Matches 144; Conservative 89; Mismatches 261; Indels 242; Gaps 31;

```

QY 30 NAEYGNSSV---FLENSTFDERGHSINDYSISPDGOFILLEYVNVYKWRHSYASDYID 86
Db 248 NAYYEMSGLEVSEELRTFG--GHDAKFIDSQENEFRLYYNKFK-----DIAS 295
QY 87 -LNKROLITEERIPNNTQWVTWSPVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIY 145
Db 296 TLNKAQIV-----GTTASLOYMKNVPEKYLLEDTSKFSVDLKF-----DKLY 342
QY 146 NGITDWMVEEVFSAYSAWLNWSPNGTFLAY--AQFNDTEVPLIEYFYSDESLOYPKTVR 203
Db 343 KMLTE-IYTEDNFVKPFKV--LNRKTYLNFDAVKINIVPKVNYTYIDGFNLRLNTNL-- 397
QY 204 VYPKAGAVNPTVKPFVNTDSSLSSVTNATSQITAPASMLIGDHY--LCDVTWATERI 261
Db 398 -----AANFNGQNTENNMNFTLKNFTG-----LFEFYKLLCVRGIITSKTK 440
QY 262 SLQWLRRIQNVSMIDICDYDESSGRMNCLVARQHIEMSTTGWGRFRPSPHFTLDGNSF 321
Db 441 SLDDDDKGYNKALNDLC-----IKYNNMDLFFSPSEDNFTDLNKG 481
QY 322 YKIISN---EEGVRHCYFOIDKKOCTFITKGTWVIGIEALTSYDL--YYISNEYKMG 376
Db 482 EETSDTNTAEAAENISLDLIQYVLTFFNDEPENISIENTSSDIIGOLELMPNIERFP 541
QY 377 GERNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAYQYLCRSGGLPLPYTLHSSV 436
Db 542 NGK---KYELDKYTFPHYL-----RAQEFHKGSRIA-----LTNSV 575
QY 437 NDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWYQWILPPHFDKSKYPLLDV 496
Db 576 NEALLN-----PSRVYTFP----- 589
QY 497 YAGPCSQKADTVFRNLNWT-----YLASTENIIVASPDGRSGVYQGDKIMH----- 542
Db 590 -----SSDYVKVKNKATEAAMFLGWVEQLVYDFDTESEVSTTDKIADITIIIPYIGP 642
QY 543 AJNRRLTGFEVEDQIEAARQFS-----KMGFVDNKRJAING-WSYGYVTSWVLGSGGVF 597
Db 643 ALN--IGNMLYKDDFVGALIFSGAVILLEFIPALIPVLGTFAVLSYIANKVLTQV--- 697
QY 598 KCGIAPVPSVRWEYDVSVYTERYMGILPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTA 657
Db 698 ---IDNLSKRNEKWDVY-----KIVITNWL 722
QY 658 DDNVH-----FQSAQISKALVDVGVDVFOAMWYTDHGH-----IASSTAHO 699
Db 723 KVNTQIDILIRKKMKEALENOAETRAI-----INYQVNYTEBEKKNINFINIDLSSKLINE 778
QY 700 HI---YTHSMHFKQC 712
Db 779 SINKAMININKFLNQC 794

```

RESULT 14
 US-11-077-550-155

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2006, 14:06:40 ; Search time 163 Seconds
(without alignments)

3099.130 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877

Sequence: 1 NTVRLKLYSLRWISDHEYL.....AQHIYTHMSHFIRKQCFSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	766	1 DPP4_HUMAN	P27487 h dipeptidyl
2	3877	100.0	766	2 Q53YH1_HUMAN	Q53tn1 homo sapien
3	3848	99.3	765	2 Q5R7G7_PONPY	Q5r7g7 pongo pygma
4	3531	91.1	765	1 DPP4_BOVIN	P81425 b dipeptidyl
5	3501	90.3	765	1 DPP4_FELCA	Q9n217 f dipeptidyl
6	3489	90.0	766	1 DPP4_PIG	P22411 s dipeptidyl
7	3354.5	86.5	767	1 DPP4_RAT	P14740 r dipeptidyl
8	3340	86.1	760	1 DPP4_MOUSE	P28843 m dipeptidyl
9	2490.5	64.2	751	2 Q758B3_AGKHB	Q758b3 agkistrodon
10	2489.5	64.2	751	2 Q758B2_AGKHB	Q758b2 agkistrodon
11	2478	63.9	759	2 Q5Z181_CHICK	Q5z181 gallus gall
12	2309	59.6	737	2 Q641D6_XENLA	Q641d6 xenopus lae
13	2292.5	69.1	748	2 P70092_XENLA	P70092 xenopus lae
14	2145	55.3	760	2 Q53TP5_HUMAN	Q53tp5 homo sapien
15	2141.5	55.2	761	2 Q8R492_RAT	Q8r492 rattus norv
16	2138	55.1	760	1 SEPR_HUMAN	Q12884 rattus norv
17	2136.5	55.1	761	1 SEPR_MOUSE	P97331 mus musculus
18	2103	54.2	755	2 Q91651_XENLA	Q91651 xenopus lae
19	1966.5	50.7	860	2 Q4RQJ4_TETNG	Q4rqj4 tetraodon n
20	1819.5	46.9	704	2 Q4S309_TETNG	Q4s309 tetraodon n
21	1215	31.4	796	1 DPP10_HUMAN	Q8n608 homo sapien
22	1207	31.1	796	1 DPP10_RAT	Q8q629 rattus norv
23	1192.5	30.8	796	2 Q6INB7_XENLA	Q6inb7 xenopus lae
24	1182	30.5	797	1 DPP10_MOUSE	Q6nxx7 mus musculus
25	1117	28.8	865	1 DPP6_HUMAN	P42658 homo sapien
26	1111.5	28.7	731	2 Q7P9F9_ANOGA	Q7p9f9 anopheles g
27	1109	28.6	803	1 DPP6_PANTR	Q5i650 pan troglod
28	1106	28.5	803	2 Q9QV78_SMURI	Q9qv78 rattus sp.
29	1104	28.5	859	1 DPP6_RAT	P46101 rattus norv
30	1101	28.4	745	2 Q9VVM2_DROME	Q9vnm2 drosophila
31	1094	28.2	863	1 DPP6_BOVIN	P42659 bos taurus

32	1077	27.8	801	2 Q5U4C2_MOUSE	Q5u4c2 mus musculus
33	1077	27.8	803	2 Q80VM5_MOUSE	Q80vm5 mus musculus
34	1077	27.8	804	1 DPP6_MOUSE	Q92218 mus musculus
35	1057.5	27.3	815	2 Q4RK63_TETNG	Q4rk63 tetraodon n
36	1052	27.1	724	2 Q5I895_9PRIM	Q5i895 saimiri bol
37	1051	27.1	604	2 Q538S4_HUMAN	Q538s4 homo sapien
38	1047.5	27.0	439	2 Q6PG25_XENLA	Q6pg25 xenopus lae
39	1045	27.0	801	2 Q4RK66_TETNG	Q4rk66 tetraodon n
40	1030	26.6	802	2 Q9VUH1_DROME	Q9vuh1 drosophila
41	1025.5	26.5	935	2 Q9VMB4_DROME	Q9vmb4 drosophila
42	998	25.7	775	2 Q5J6J3_TIRIRU	Q5j6j3 trichophyto
43	996	25.7	842	2 Q4RVF8_TETNG	Q4rvf8 tetraodon n
44	975	25.1	874	2 Q7SI80_EMENI	Q7si80 emericella
45	967	24.9	880	2 Q5B934_EMENI	Q5b934 aspergillus

ALIGNMENTS

RESULT 1
DPP4_HUMAN
ID DPP4_HUMAN STANDARD; PRT; 766 AA.
AC P27487;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE IV) (T-cell activation antigen CD26) (TP103) (Adenosine deaminase
DE complexing protein 2) (ADABP) [Contains: Dipeptidyl peptidase 4
DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl
DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].
GN Name=DPP4; Synonym=ADCP2, CD26;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=92329551; PubMed=1352704; DOI=10.1016/0167-4781(92)90036-Y;
RA Misumi Y., Hayashi Y., Arakawa F., Ikehara Y.;
RT "Molecular cloning and sequence analysis of human dipeptidyl peptidase
RL IV, a serine proteinase on the cell surface.";
RL Biochim. Biophys. Acta 1131:333-336(1992).
[2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=95012454; PubMed=7927537;
RA Abbott C.A., Baker E., Sutherland G.R., McCaughan G.W.;
RT "Genomic organization, exact localization, and tissue expression of
RL the human CD26 (dipeptidyl peptidase IV) gene.";
RL Immunogenetics 40:331-338(1994).
[3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=peripheral blood;
RX MEDLINE=92325476; PubMed=1352530;
RA Tanaka T., Camerini D., Seed B., Torimoto Y., Dang N.H., Kameoka J.,
RA Dahlberg H.N., Schlossman S.F., Morimoto C.;
RT "Cloning and functional expression of the T cell activation antigen
RL CD26.";
RL J. Immunol. 149:481-486(1992).
[4]
RP ERRATUM.
RX MEDLINE=93171637; PubMed=8094732;
RA Tanaka T.;
RL J. Immunol. 150:2090-2090(1993).
[5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=prostate, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klauener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[6]
RN NUCLEOTIDE SEQUENCE OF 1-551.
RP TISSUE=Colon;
RX MEDLINE=92165847; PubMed=1347043;
RA Darmoul D., Lacasa M., Baricault L., Marguet D., Sapin C., Trotot P.,
RA Barbat A.;
RT "Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like
RT colon cancer cell lines HT-29 and Caco-2. Cloning of the complete
RT human coding sequence and changes of dipeptidyl peptidase IV mRNA
RT levels during cell differentiation.";
RL J. Biol. Chem. 267:4824-4833(1992).
[7]
RN NUCLEOTIDE SEQUENCE OF 545-766.
RP TISSUE=Colon;
RX MEDLINE=91024044; PubMed=1977364;
RA Darmoul D., Lacasa M., Chantret I., Swallow D., Trugnan G.;
RT "Isolation of a cDNA probe for the human intestinal
RT dipeptidylpeptidase IV and assignment of the gene locus DPP4 to
RT chromosome 2.";
RL Ann. Hum. Genet. 54:191-197(1990).
[8]
RN NUCLEOTIDE SEQUENCE OF 1-31.
RX MEDLINE=96067599; PubMed=7487939;
RA Boehm S.K., Gum J.R. Jr., Erickson R.H., Hicks J.W., Kim Y.S.;
RT "Human dipeptidyl peptidase IV gene promoter: tissue-specific
RT regulation from a TATA-less GC-rich sequence characteristic of a
RT housekeeping gene promoter.";
RL Biochem. J. 311:835-843(1995).
[9]
RN PROTEIN SEQUENCE OF 1-22, AND TISSUE SPECIFICITY.
RP PubMed=1677636;
RA Gorvel J.P., Ferrero A., Chambraud L., Rigal A., Bonicel J.,
RA Maroux S.;
RT "Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human
RT small intestine and colon.";
RL Gastroenterology 101:618-625(1991).
[10]
RN PARTIAL PROTEIN SEQUENCE.
RP TISSUE=Kidney;
RX MEDLINE=93210468; PubMed=8096237; DOI=10.1084/jem.177.4.1135;
RA Morrison M.E., Vijayasekaradi S., Engelstein D., Albino A.P.,
RA Houghton A.N.;
RT "A marker for neoplastic progression of human melanocytes is a cell
RT surface ectopeptidase.";
RL J. Exp. Med. 177:1135-1143(1993).
[11]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 38-766 IN COMPLEX WITH
RP INHIBITOR, AND HOMODIMERIZATION
RX PubMed=12832764; DOI=10.1107/S0907444903010059;
RA Oefner C., D'Arcy A., Mac Sweeney A., Pierau S., Gardiner R.,
RA Dale G.E.;
RT "High-resolution structure of human apo dipeptidyl peptidase IV/CD26
RT and its complex with 1-[(2-{[5-iodopentylidene-2-yl]amino}-ethyl)amino]-
RT acetyl]-2-cyano-(S)-pyrrolidine.";
RL Acta Crystallogr. D 59:1206-1212(2003).
[12]
RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-771, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12646248; DOI=10.1016/S0006-291X(03)00258-4;
RA Hiramatsu H., Inoko K., Higashiyama Y., Fukushima C., Shima H.,
RA Sugiyama S., Yano K., Yamamoto A., Shimizu R.;
RT "The structure and function of human dipeptidyl peptidase IV,
RT possessing a unique eight-bladed beta-propeller fold.";
RL Biochem. Biophys. Res. Commun. 302:849-854(2003).
[13]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12483204; DOI=10.1038/nsb882;
RA Rasmussen H.B., Branner S., Wiberg F.C., Wagtmann N.;
RT "Crystal structure of human dipeptidyl peptidase IV/CD26 in complex
RT with a substrate analog.";
RL Nat. Struct. Biol. 10:19-25(2003).
[14]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12906826; DOI=10.1016/S0969-2126(03)00160-6;
RA Thoma R., Loeffler B., Stihle M., Huber W., Ruf A., Hennig M.;
RT "Structural basis of proline-specific exopeptidase activity as
RT observed in human dipeptidyl peptidase-IV.";
RL Structure 11:947-959(2003).
CC -!- FUNCTION: Removes N-terminal dipeptides sequentially from the
CC polypeptides having unsubstituted N-termini provided that the
CC penultimate residue is proline. Plays a role in T cell activation.
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -!- SUBUNIT: Homodimer or heterodimer with seprase (FAP).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC soluble form.
CC -!- TISSUE SPECIFICITY: Expressed in the poorly differentiated crypt
CC cells of the small intestine as well as in the mature villous
CC cells. Expressed at very low levels in the colon.
CC -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC by proteolytic processing.
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; U13735; AAB60646.1; Genomic DNA.
CC EMBL; U13710; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13711; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13712; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13713; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13714; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13715; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13716; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13717; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13718; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13719; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13720; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13721; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13722; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13723; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13724; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13725; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13726; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13727; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13728; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13729; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13730; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13731; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13732; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13733; AAB60646.1; JOINED; Genomic DNA.
CC EMBL; U13734; AAB60646.1; JOINED; Genomic DNA.

```
DR EMBL; M74777; AAA51943.1; -; mRNA.
DR EMBL; BC013329; AAH13329.2; -; mRNA.

Query Match      100.0%; Score 3877; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 3.6e-256;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSITDFDEFGHSINDYSISPD 60
Db 51 NTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSITDFDEFGHSINDYSISPD 110
Qy 61 GQFILLENNYKQWRHSYASDYIDLKRLQILTEERIPNNTQWTVSPVGHKLAYVWNN 120
Db 111 GQFILLENNYKQWRHSYASDYIDLKRLQILTEERIPNNTQWTVSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 230
Qy 181 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290
Qy 241 ASMLIGDHYLDCVWTWATQERISLQWLRRIQNYSDICDYDESSGRWNCVLAHQHIE 300
Db 291 ASMLIGDHYLDCVWTWATQERISLQWLRRIQNYSDICDYDESSGRWNCVLAHQHIE 350
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFITKGTWVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFITKGTWVIGIEAL 410
Qy 361 TSDYLYYISNEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDYLYYISNEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470
Qy 421 RCGSPGLPLTYLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFILNETKFWYQ 480
Db 471 RCGSPGLPLTYLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFILNETKFWYQ 530
Qy 481 PPHFDKSKKYPILLDVTAGCSQKADTVFRLNWTATYLASTENIIVASFDFGRSGYQ 540
Db 531 PPHFDKSKKYPILLDVTAGCSQKADTVFRLNWTATYLASTENIIVASFDFGRSGYQ 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNKRIAIWGSYGGYVTVSMVLSGSGVFK 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVNKRIAIWGSYGGYVTVSMVLSGSGVFK 650
Qy 601 IAVAPSRWEYSDVYTERYMGVLTPEPDLNHYRNSTVMSRAENFKQVEYLLIHGTADD 660
Db 651 IAVAPSRWEYSDVYTERYMGVLTPEPDLNHYRNSTVMSRAENFKQVEYLLIHGTADD 710
Qy 661 VHFQSSAQISKALVDGVDFQAMWYTTDEHDGIIASSTAHOIYTHMSHFIKQCFSLP 716
Db 711 VHFQSSAQISKALVDGVDFQAMWYTTDEHDGIIASSTAHOIYTHMSHFIKQCFSLP 766
```

RESULT 2

```
Q53TN1 HUMAN
ID Q53TN1_HUMAN PRELIMINARY; PRT; 766 AA.
AC Q53TN1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein DPP4.
GN Name=DPP4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ozanich A., Stoneking T., Hawkins M., Sapetti L.,
```

```
RT "The sequence of Homo sapiens BAC clone RP11-178A14.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RA NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RA NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RA NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC008063; AAX93179.1; -; Genomic_DNA.
DR SMR; O53TN1; 39-766.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
DR HydroLase; Hypothetical protein; Transmembrane.
SQ SEQUENCE 766 AA; 88279 MW; 5FB4A2C6662D6117 CRC64;
```

```
Query Match      100.0%; Score 3877; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 3.6e-256;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 NTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSITDFDEFGHSINDYSISPD 60
Db 51 NTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSITDFDEFGHSINDYSISPD 110
Qy 61 GQFILLENNYKQWRHSYASDYIDLKRLQILTEERIPNNTQWTVSPVGHKLAYVWNN 120
Db 111 GQFILLENNYKQWRHSYASDYIDLKRLQILTEERIPNNTQWTVSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 230
Qy 181 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290
Qy 241 ASMLIGDHYLDCVWTWATQERISLQWLRRIQNYSDICDYDESSGRWNCVLAHQHIE 300
Db 291 ASMLIGDHYLDCVWTWATQERISLQWLRRIQNYSDICDYDESSGRWNCVLAHQHIE 350
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFITKGTWVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFITKGTWVIGIEAL 410
Qy 361 TSDYLYYISNEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDYLYYISNEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470
Qy 421 RCGSPGLPLTYLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFILNETKFWYQ 480
Db 471 RCGSPGLPLTYLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFILNETKFWYQ 530
Qy 481 PPHFDKSKKYPILLDVTAGCSQKADTVFRLNWTATYLASTENIIVASFDFGRSGYQ 540
Db 531 PPHFDKSKKYPILLDVTAGCSQKADTVFRLNWTATYLASTENIIVASFDFGRSGYQ 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNKRIAIWGSYGGYVTVSMVLSGSGVFK 600
```

```
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGYVTSMVLGSGGVFKCG 650
QY 601 IAVAPVSRWEYDSVYTERYMGFLPTEDNLDHYRNSTVMSRAENFQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDSVYTERYMGFLPTEDNLDHYRNSTVMSRAENFQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDGVDFQAMWYTDDEHGHASSTAHOHIYTHMSHFHKQCFSLP 716
Db 711 VHFQSAQISKALVDGVDFQAMWYTDDEHGHASSTAHOHIYTHMSHFHKQCFSLP 766

RESULT 3
QSR7G7_PONPY
ID QSR7G7_PONPY PRELIMINARY; PRT; 765 AA.
AC
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKF2p469P1419.
GN Name=DKF2p469P1419;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Kidney;
RG The German cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR860150; CH92293.1; -; mRNA.
DR SRR; QSR7G7; 38-765.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept S9 AS.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase S9B.
DR InterPro; IPR000379; Ser. esterase.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase S9; 1.
DR PROSITE; PS00708; PRO ENDOPEP SER; 1.
DR Hydrolase; Hypothetical protein; Transmembrane.
KW SEQUENCE 765 AA; 88038 MW; 6F81ECD98C3D2397 CRC64;

Query Match 99.3%; Score 3848; DB 2; Length 765;
Best Local Similarity 99.2%; Pred. No. 3.5e-254;
Matches 710; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NTVRLKYLRLWTSDEHYLYKQNNILVFNAEYGNSSVFLENSTPDEFGHSINDYSISPD 60
Db 50 NTVRLKYLRLWTSDEHYLYKQNNILVFNAEYGNSSAFLENSTPDEFGHSINDYSISPD 109
QY 61 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEIRIPNNTQWTVSPVGHKLAYVWNN 120
Db 110 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEIRIPNNTQWTVSPVGHKLAYVWNN 169
QY 121 DIYVKLEPNLPSRIWTGKEDIYNGITDWYEEVFSAYSALWWSPNGTFLAYAQFND 180
Db 170 DIYVKLEPNLPSRIWTGKEDIYNGITDWYEEVFSAYSALWWSPNGTFLAYAQFND 229
QY 181 TEVPLEYGFYSDLSQYKPTVVRYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 240
Db 230 TEVPLEYGFYSDLSQYKPTVVRYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 289
QY 241 ASMLIGDHVLCVDTWATQERISLQWLRRIQNTYSVMDICDYDESSGRWNCVLVARQHIEMST 300
Db 290 ASMLIGDHVLCVDTWATQERISLQWLRRIQNTYSVMDICDYDESSGRWNCVLVARQHIEMST 349
```

```
QY 301 TGMVGRFRSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTPITTKGTWEVIGIAL 360
Db 350 TGMVGRFRSEPHFTSDGNSFYKLIISNEEGYRHCYFQIDKKDCTPITTKGTWEVIGIAL 409
QY 361 TSDYLYIYSNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 410 TSDYLYIYSNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 469
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQML 480
Db 470 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQML 529
QY 481 PPHFDKSKYPLLLDVIYAGCSQKADTVFRLNWTATLASTENIIVASFGRSGSYQGDKI 540
Db 530 PPHFDKSKYPLLLDVIYAGCSQKADTVFRLNWTATLASTENIIVASFGRSGSYQGDKI 589
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGYVTSMVLGSGGVFKCG 600
Db 590 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIAIWGWSYGYVTSMVLGSGGVFKCG 649
QY 601 IAVAPVSRWEYDSVYTERYMGFLPTEDNLDHYRNSTVMSRAENFQVEYLLIHGTADDN 660
Db 650 IAVAPVSRWEYDSVYTERYMGFLPTEDNLDHYRNSTVMSRAENFQVEYLLIHGTADDN 709
QY 661 VHFQSAQISKALVDGVDFQAMWYTDDEHGHASSTAHOHIYTHMSHFHKQCFSLP 716
Db 710 VHFQSAQISKALVDGVDFQAMWYTDDEHGHASSTAHOHIYTHMSHFHKQCFSLP 765

RESULT 4
DPP4_BOVIN STANDARD; PRT; 765 AA.
ID DPP4_BOVIN
AC P81425; Q8WVG8;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE IV) (T-cell activation antigen CD26) (Adenosine deaminase complexing
DE protein) (ADCP-I) (Activation molecule 3) (ACT3) (WC10) [Contains:
DE Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane
DE form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV
DE soluble form)].
GN Name=DPP4; Synonyms=CD26;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Lymphocyte;
RX MEDLINE=22067734; PubMed=12073152; DOI=10.1007/s00251-002-0456-6;
RA Lee S.-U., Park Y.-H., Davis W.C., Hamilton M.J., Naessens J.,
RA Bohach G.A.;
RT "Molecular characterization of bovine CD26 upregulated by a
RT staphylococcal superantigen.";
RL Immunogenetics 54:216-220(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-17.
RC Tissue=Thymus;
RX MEDLINE=20211197; PubMed=11981836;
RX DOI=10.1007/s121-4141(200205)32:5<1472::AID-IMMU1472>3.0.CO;2-Q;
RA Gliddon D.R., Howard C.;
RT "CD26 is expressed on a restricted subpopulation of dendritic cells in
RT vivo.";
RL Eur. J. Immunol. 32:1472-1481(2002).
RN [3]
RP PROTEIN SEQUENCE OF 1-24.
RC Tissue=T-cell;
RX MEDLINE=21482004; PubMed=11598101;
RX DOI=10.1128/IAI.69.11.7190-7193.2001;
RA Lee S.-U., Ferens W., Davis W.C., Hamilton M.J., Park Y.-H., Fox L.K.,
```


RA Naessens J., Bobach G.A.;
 RT "Identity of activation molecule 3 on superantigen-stimulated bovine
 RL cells is CD26.";
 RN Infect. Immun. 69:7190-7193(2001).
 RN [4]
 RP PROTEIN SEQUENCE OF 537-546.
 RC TISSUE=Kidney;
 RX MEDLINE=98293306; PubMed=9629661; DOI=10.1016/S0305-0491(97)00327-1;
 RA Ben-Shooshan I., Parola A.H.;
 RT "The Cp-I subunit of adenosine deaminase complexing protein from calf
 RT kidney is identical to human, mouse, and rat dipeptidyl peptidase
 RT IV.";
 RL Comp. Biochem. Physiol. 119B:289-292(1998).
 CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from
 CC polypeptides having unsubstituted N-termini provided that the
 CC penultimate residue is proline. Binds and regulates the activity
 CC of ADA.
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-l-
 CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
 CC Zaa is neither Pro nor hydroxyproline.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
 CC soluble form (By similarity).
 CC -1- TISSUE SPECIFICITY: Intestinal epithelium, dendritic cells and
 CC several immune system tissues.
 CC -1- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
 CC by proteolytic processing (By similarity).
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; AF461806; AAL67836.1; -; mRNA.
 DR EMBL; AY056834; AAL23628.1; -; mRNA.
 DR HSSP; P27487; 1PFP.
 DR SMR; P81425; 38-764.
 DR MEROPS; S09.003; -.
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser esters.
 DR Pfam; PF00930; DPPIV N1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;
 KW Protease; Serine protease; Signal-anchor; Transmembrane.
 FT CHAIN 1 765
 FT Dipeptidyl peptidase 4 soluble form (By
 FT similarity).
 FT TOPO_DOM 1 6
 FT Signal-anchor for type II membrane
 FT TRANSMEM 7 29
 FT protein (Potential).
 FT Extracellular (Potential).
 FT ACT_SITE 30 765
 FT Charge relay system (By similarity).
 FT ACT_SITE 629 629
 FT Charge relay system (By similarity).
 FT ACT_SITE 707 707
 FT Charge relay system (By similarity).
 FT CARBOHYD 739 739
 FT N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 84 84
 FT N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 91 91
 FT N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 149 149
 FT N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 218 218
 FT N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 228 228
 FT N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 271 271
 FT N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 280 280
 FT N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 320 320
 FT N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 392 392
 FT N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 495 495
 FT N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 684 684
 FT By similarity.
 FT DISULFID 384 393
 FT By similarity.
 FT DISULFID 443 446
 FT By similarity.
 FT DISULFID 453 471

FT DISULFID 648 761 By similarity.
 SQ SEQUENCE 765 AA; 88369 MW; E32165421F43E116 CRC64;
 Query Match 91.1%; Score 3531; DB 1; Length 765;
 Best Local Similarity 89.4%; Pred. No. 1.6e-232;
 Matches 639; Conservative 40; Mismatches 36; Indels 0; Gaps 0;
 QY 1 NTYRLKLYSLRWISDHEYLKQENNILVFNAAEYGNSSVFLNENSTFDFEGHSHINDYSISPD 60
 DB 50 NTFRMKFYNLRWSDHEYLKQENNILVFNAAEYGNSSVFLNENSTFDFEGHSHINDYSISPD 109
 QY 61 GQFILLVYVVKWRHSYASDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 120
 DB 110 RQVILFEYVVKWRHSYASDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 169
 QY 121 DIVVKTEPNLPSVRIITWTKGEDIYNGITDWVVEEVEFSAYLSALWSPNCTFLAYAQFND 180
 DB 170 DIIVKNEPNSQRITWTKGVYINGITDWVVEEVEFSAYLSALWSPNCTFLAYAQFND 229
 QY 181 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 240
 DB 230 TEVPLIEYSFYSDLSQYPKTVKIPYPKAGAVNPTKFFVNNISLSPINATSIQIVPP 289
 QY 241 ASMLIGDHLYCDVTWATQERISLQWRRIQNTSYVMDICDYDESSGRNCLVARQHTEMST 300
 DB 290 GSVLIGDHLYCDVTWVTEERISLQWRRIQNTSYVMDICDYDRSTGRWISSVGRQHIEIST 349
 QY 301 TGWVGFRSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKDKCTFITKGTWVIGIEAL 360
 DB 350 TGWVGFRPAEPHFTSDGNSFYKIIISNEGYRHCIFQTDKRNCTFITKGAWEVIGIEAL 409
 QY 361 TSDYLYYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
 DB 410 TSDYLYYIISNEYKMGPGARNLYKIQLNDYTKVTCLSCELNPERCQYYSVSFSQEAKEYQL 469
 QY 421 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETXFTWOMIL 480
 DB 470 RCGPGGLPLYTLHNSNNDKELRVLENNSDLDQVLDQVQMPSSKKLDPIILHGTFTFYQMI 529
 QY 481 PPHEDKSKYPLLLDVYAGCSQKADTVFLNWNATYLASTENIIVASFQGRSGSGYQGDKI 540
 DB 530 PPHEDKSKYPLLLLEVYAGCSQKADAIPLNWNATYLASTENIIVASFQGRSGSGYQGDKI 589
 QY 541 MEAINRRLGTFEVDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSMLGSGSVFKKCG 600
 DB 590 MEAINRRLGTFEVDQIEATRFQSKMGFVNDKRIAIWGSYGGYVTSMLGSGSVFKKCG 649
 QY 601 IAVAPVSRWEYSDSVYTERYMGFLPTPEDNLNHYRSTVMSRAENFKQVEYLLIHGTADDN 660
 DB 650 IAVAPVSKWEYSDSVYTERYMGFLPTPEDNLNHYRSTVMSRAENFKQVEYLLIHGTADDN 709
 QY 661 VHFQSAQISKALVDVGVDFQAMWYTDHGHGASSTAHQHIYTHMSHFHKQCFSL 715
 DB 710 VHFQSAQISKALVDVGVDFQSMWYTDHGHGASSTAHQHIYTHMSHFHKQCFSL 764

RESULT 5

DPP4_FELCA

ID DPP4_FELCA STANDARD; PRT; 765 AA.

AC Q9N2I7;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP

DE IV) (T-cell activation antigen CD26) [Contains: Dipeptidyl peptidase 4

DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl

DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].

GN Name=DPP4; Synonyms=CD26;

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;

OC Felinae; Felis.

OX NCBI_TaxID=9685;

RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Peripheral blood;
 RX MEDLINE=20094000; PubMed=10630304; DOI=10.1007/s002510050616;
 RA Nishimura Y., Miyazawa T., Ikeda Y., Izumiya Y., Nakamura K., Sato E.,
 RA Mikami T., Takahashi E.;
 RT "Molecular cloning and sequencing of a cDNA encoding the feline T-cell
 RT activation cloning and sequencing of a cDNA encoding the feline T-cell
 RL Immunogenetics 50:366-368(1999).
 CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from
 CC polypeptides having unsubstituted N-termini provided that the
 CC penultimate residue is proline (By similarity).
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
 CC Zaa- from a polypeptide, preferentially when Yaa is Pro, provided
 CC Zaa is neither Pro nor hydroxyproline.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
 CC soluble form (By similarity).
 CC -1- PTM: The soluble form (SDPP) derives from the membrane form (WDPP)
 CC by proteolytic processing (By similarity).
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AB023952; BRA92344.1; -; mRNA.
 DR HSSP; P27487; 1PFQ.
 DR SMR; Q9N217; 38-765.
 DR MEROPS; S09.003; -;
 DR InterPro; IPR002471; Pept S9_AS.
 DR InterPro; IPR001375; Peptidase S9.
 DR InterPro; IPR002469; Peptidase S9B.
 DR InterPro; IPR000379; Ser esters.
 DR Pfam; PF00930; DPPIV_N; 1.
 DR Pfam; PF00326; Peptidase S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Aminopeptidase; Glycoprotein; Hydrolase; Protease; Serine protease;
 KW Signal-anchor; Transmembrane.
 FT CHAIN 1 765 Dipeptidyl peptidase 4 membrane form.
 FT CHAIN 38 765 Dipeptidyl peptidase 4 soluble form (By
 FT similarity).
 FT TOPO_DOM 1 6 Cytoplasmic (Potential).
 FT TRANSMEM 7 29 Signal-anchor for type II membrane
 FT protein (Potential).
 FT TOPO_DOM 30 765 Extracellular (Potential).
 FT ACT_SITE 629 629 Charge relay system (By similarity).
 FT ACT_SITE 707 707 Charge relay system (By similarity).
 FT ACT_SITE 739 739 Charge relay system (By similarity).
 FT CARBOHYD 84 84 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 91 91 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 149 149 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 178 178 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 228 228 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 280 280 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 320 320 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 330 330 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 331 331 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 519 519 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 684 684 N-linked (GlcNAc...) (By similarity).
 FT DISULFID 384 393 By similarity.
 FT DISULFID 443 446 By similarity.
 FT DISULFID 453 471 By similarity.
 FT DISULFID 648 761 By similarity.
 SQ SEQUENCE 765 AA; 88213 MW; 3EPCE98A22B175D9 CRC64;
 Query Match 90.3%; Score 3501; DB 1; Length 765;
 Best Local Similarity 88.0%; Pred. No. 1.8e-230;
 Matches 630; Conservative 46; Mismatches 40; Indels 0; Gaps 0;
 1 NTYRLKXLSLWISDHELYKQENNLVFNFAEYGNSSVFLFNSTPDEFGHSINDYSISPD 60

Db 50 NTFRVKFYSLRWVSDHDYLYKQDNILLFNAEYGNSSVFLFNSTPDEFGHSINDYSISPD 109
 QY 61 GQFILLENNYVQWRHSYTSASVDIYDLNKRQLITERRIPNNTQMTVTSVGVKHLAYVWN 120
 Db 110 GQFILLENNYVQWRHSYTSASVDIYDLNKRQLITBEKIPNNTQMTVTSVGVKHLAYVWN 169
 QY 121 DIYVKIENLPSYRITWTGKEDIYNGITDWTVEEVSAYSALWMSPNGTFLAYAQFND 180
 Db 170 DYYVKNENPSSSHRITWTGEEINAYNGIADWYBEEIFSAYSALWMSPGTFLAYAQFND 229
 QY 181 TEVPLIEYSFYDESLSQYPKTVRVYPKAGAVNPTVKFFVNNTDSLSSTVNTATSIQITAP 240
 Db 230 TQVPLIEYSFYDESLSQYPMWNRIPYPKAGANPTVKLVKITDNLNPNNTATSEITPP 289
 QY 241 ASMLIGDHYLCVDTWATERISQWLRLRIQNTYSVMDICDYDESSGRNCLVARQHIEMST 300
 Db 290 AAMLGDDYLCVDTWANEERISQWLRLRIQNTYSVMDIRDYNNSTGKWISSAAQEHIEEMST 349
 QY 301 TGMVGRFRPSEPHTLDGNSFYKIIISNEEGVHICYFOIDKDKDCTFTTKGTWEGVIGIAL 360
 Db 350 TGMVGRFRPAEPHTSDGRNFYKIIISNEEGYKHCIFQIDKDKDCTFTTKGTWEGVIGIAL 409
 QY 361 TSDYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOVYSFSKEAKYQOL 420
 Db 410 TTDYLYISNEYKMGPGGRNLYKIQLDYTKVACLSCELKPERCOVYSFSKEAKYQOL 469
 QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLPDIFILNETKFWYQML 480
 Db 470 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQEVQMPSSKKLPDIFILNETKFWYQML 529
 QY 481 PHFPDKSKYPLLLDYAGPCSQKADTVFRLNWTYLASTENIIVASFDGRSGYQGDKI 540
 Db 530 PHFPDTSKKYPLLLDYAGPCSQKADAFRLNWTYLASTENIIVASFDGRSGYQGDKI 589
 QY 541 MHAINRRLGTFRVEDQIEAARQFSKMGFVDNKRKRIAGWSYGGYVTSMLVSGSGYVFKCG 600
 Db 590 MHAVNRRLGTFRVEDQIEAARQFSKMGFVDNKRKRIAGWSYGGYVTSMLVSGSGYVFKCG 649
 QY 601 IAVAPVSRWEYYSVYTERYMGSLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
 Db 650 IAVAPVSRWEYYSVYTERYMGSLPTPDNLDYKNTVMSRAENFKQVEYLLIHGTADDN 709
 QY 661 VHFQSAQISKALVDGVDFQAMWYTDHGHGTASTAHQHIYTHSHFHKQCFSLP 716
 Db 710 VHFQSAQISKALVDAGVDFQAMWYTDHGHGTASTAHQHIYTHSHFHKQCFSLP 765
 RESULT 6
 DPP4_PIG
 ID DPP4_PIG STANDARD; PRT; 766 AA.
 AC P22411; Q866G3;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
 DE IV) (T-cell activation antigen CD26) [Contains: Dipeptidyl peptidase 4
 DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl
 DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].
 GN Name=DPP4; Synonyms=CD26;
 OS Sus scrofa (Pig)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
 OC Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP NUCLEOTIDE SEQUENCE, AND FUNCTION.
 RC TISSUE=Kidney;
 RX PubMed=14719797; DOI=10.1515/BC.2003.172;
 RA Baer J., Weber A., Hoffmann T., Stork J., Wermann M., Wagner L.,
 RA Aust S., Gehartz B., Demuth H.-U.;
 RT "Characterisation of human dipeptidyl peptidase IV expressed in Pichia
 RT pastoris. A structural and mechanistic comparison between the

recombinant human and the purified porcine enzyme.";
 [2]
 RN Biol. Chem. 384:1553-1563(2003).
 RP NUCLEOTIDE SEQUENCE OF 2-67.
 RC TISSUE=Kidney;
 RX MEDLINE=94093209; PubMed=7903569;
 RA Thomsen P.D., Qvist H., Marklund L., Andersson L., Sjostrom H.,
 Noren O.;
 RA "Assignment of the dipeptidylpeptidase IV (DPP4) gene to pig
 chromosome 15q21.";
 RT chromsome 15q21.";
 RL Mamm. Genome 4:604-607(1993).
 RN [3]
 RP PROTEIN SEQUENCE OF 38-71.
 RC TISSUE=Kidney;
 RX MEDLINE=91273813; PubMed=1675855;
 RA Seidl R., Mann K., Schaeffer W.;
 RA "N-terminal amino-acid sequence of pig kidney dipeptidyl peptidase IV
 solubilized by autolysis.";
 RT Biol. Chem. Hoppe-Seyler 372:213-214(1991).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 39-766, N-GLYCOSYLATION
 SITES, AND HOMODIMERIZATION.
 RX PubMed=12690074; DOI=10.1073/pnas.0230620100;
 RA Engel M., Hoffmann T., Wagner L., Wermann M., Heiser U.,
 Kieffersauer R., Huber R., Bode W., Demuth H.-U., Brandstetter H.;
 RT "The crystal structure of dipeptidyl peptidase IV (CD26) reveals its
 functional regulation and enzymatic mechanism.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:5063-5068(2003).
 CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from
 CC polypeptides having unsubstituted N-termini provided that the
 CC penultimate residue is proline.
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
 CC Zaa, from a polypeptide, preferentially when Yaa is Pro, provided
 CC Zaa is neither Pro nor hydroxyproline.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
 CC soluble form.
 CC -1- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
 CC by proteolytic processing.
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR ENBL; AY198323; AA043404.1; -; mRNA.
 DR ENBL; X73276; CAAS1717.1; -; mRNA.
 DR PIR; I47134; I47134.
 DR PDB; 1ORV; X-ray; A/B/C/D=39-766.
 DR PDB; 1ORW; X-ray; A/B/C/D=39-766.
 DR SNR; P22411; 39-766.
 DR MEROPS; S09.003; -.
 DR InterPro; IPR002471; Pept S9 AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_eastrs.
 DR Pfam; PF00930; DPPIV_N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_S9; 1.
 KW 3D-structure; Aminopeptidase; Direct protein sequencing; Glycoprotein;
 KW Hydrolase; Protease; Serine protease; Signal-anchor; Transmembrane.
 FT CHAIN 1 766 Dipeptidyl peptidase 4 membrane form.
 FT CHAIN 38 766 Dipeptidyl peptidase 4 soluble form.
 FT TOPO_DOM 1 6 Cytoplasmic (Potential).
 FT TRANSMEM 7 27 Signal-anchor for type II membrane
 FT protein (Potential).
 FT TOPO_DOM 28 766 Extracellular (Potential).
 FT ACT_SITE 630 766 Charge relay system (By similarity).
 FT ACT_SITE 708 708 Charge relay system (By similarity).
 FT ACT_SITE 740 740 Charge relay system (By similarity).
 FT CARBOHYD 85 85 N-linked (GlcNAc...).

FT	CARBOHYD	92	92	N-linked (GlcNAc...)	
FT	CARBOHYD	150	150	N-linked (GlcNAc...)	(By similarity).
FT	CARBOHYD	179	179	N-linked (GlcNAc...)	(Potential).
FT	CARBOHYD	219	219	N-linked (GlcNAc...)	(By similarity).
FT	CARBOHYD	229	229	N-linked (GlcNAc...)	
FT	CARBOHYD	279	279	N-linked (GlcNAc...)	
FT	CARBOHYD	321	321	N-linked (GlcNAc...)	
FT	CARBOHYD	393	393	N-linked (GlcNAc...)	(Potential).
FT	CARBOHYD	685	685	N-linked (GlcNAc...)	
FT	DISULFID	385	394		
FT	DISULFID	444	447		
FT	DISULFID	454	472		
FT	DISULFID	649	762		
FT	CONFLICT	32	32	Missing (in Ref. 2).	
SQ	SEQUENCE	766 AA;	88242 MW;	8800D520BAEA856D CRC64;	
Query Match 90.0%; Score 3489; DB 1; Length 766;					
Best Local Similarity 88.1%; Pred. No. 1.2e-229;					
Matches 631; Conservative 45; Mismatches 40; Indels 0; Gaps 0;					
QY	1	NTVRLKLSLRWISDHEYLKQENNLVFNAEYGNSSVFNLENSTFDFEGHSINDYSISPD	60		
DB	51	STFRVAFYTLQWISDHEYLKQENNLVFNAEYGNSSIFLENSTFDELGHSINDYSISPD	110		
QY	61	GQFILLVYVVKWRHSYTSYDIYDLNKRQLITEBRIPNNTQWTVTSPVGHKLAYVWNN	120		
DB	111	RQFILFEYVVKWRHSYTSYDIYDLNKRQLITEBRIPNNTQWTVTSPVGHKLAYVWNN	170		
QY	121	DIYVKLEPNLPSYRIWTGKEDIYNGITDWTVEEVFSAYSALMWSNPTFLAYAQFND	180		
DB	171	DIYVKNEPLSSQRIITWGKENVYNGVTDWVVEEVFSAYSALMWSNPTFLAYAQFND	230		
QY	181	TEVPLEYFSYDESLOYPKTVRVPYKAGVNPVTKFFVNTDLSVSNATSIQITAP	240		
DB	231	TEVPLEYFSYDESLOYPKTVRIPYKAGENPTVKFFVNTDLSVSNATSIQITAP	290		
QY	241	ASMLGDHGLCDVTWATQERISLQWLRRIQNTSVMDICDYDESGRWNCILVARQHIEMST	300		
DB	291	ASVLGDHGLCGVTWVTEERISLQWIRRAQNSIIDIDYDESGRWISSVARQHIEST	350		
QY	301	TGWVGRFRSEPHFTLDGNSFYKISNREGYRHCYFQDKDKDCTFITKGTWEVIGIEAL	360		
DB	351	TGWVGRFRSEPHFTSDGNSFYKISNREGYRHCYFQDKDKSCTFITKGTWEVIGIEAL	410		
QY	361	TSDLVYIINEYKMGPGGRNLKYIOISDYTKVTCLSCELNPERCOYYSVSFSKEAYQOL	420		
DB	411	TSDLVYIINEYKMGPGGRNLKYIQNDYTKVTCLSCELNPERCOYYSVSFSKEAYQOL	470		
QY	421	RCSGPGLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDPIILNETKFWYQML	480		
DB	471	RCFGPGLPYTLHSSSSDKELRVLEDNSALDKMLQNVQMPKSKLDVILNHLGTFKFWYQML	530		
QY	481	PHFDKSKYPLLLDVYAGCSQKADTVFRLNWTATYLASTENIIIVASFDGRGSGYQGDKI	540		
DB	531	PHFDKSKYPLLLIEVYAGCSQKVDVTFRLSWATYLASTENIIIVASFDGRGSGYQGDKI	590		
QY	541	MEAINRRLCTFEVEDOIEAARQFSKMGFVDNKRKRIATWGHYSYGYVTSMVLGSGSGYKCG	600		
DB	591	MEAINRRLCTFEVEDOIEATRFKMGFVDDKRKRIATWGHYSYGYVTSMVLGSGSGYKCG	650		
QY	601	IIVAPVSRWEYVDSVYTYMGLPTPEDNLDHRYNSTVMSRAENFKQVEYLHGTADDN	660		
DB	651	IIVAPVSKWEYVDSVYTYMGLPTPEDNLDHRYNSTVMSRAENFKQVEYLHGTADDN	710		
QY	661	VHFQSSAQISKALVDVGVDFQAWYVTDHGIASSTAHOIYTHMSHFHKQCFSLP	716		
DB	711	VHFQSSAQISKALVDAGVDFQWYVTDHGIASNAHQAHOIYTHMSHFHKQCFSLP	766		
RESULT 7					
DPP4_RAT					
ID	DPP4_RAT	STANDARD;	PRT;	767 AA.	
AC	P14740;				

DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (RC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE IV) (T-cell activation antigen CD26) (GP110 glycoprotein) (Bile
DE canalicular domain-specific membrane glycoprotein) [Contains:
DE Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane
DE form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV
DE soluble form); Dipeptidyl peptidase 4 60 kDa soluble form (Dipeptidyl
DE peptidase IV 60 kDa soluble form)].
GN Name:Dpp4; Synonyms=cd26;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RP MEDLINE=89123496; PubMed=2563382;
RX Ogata S., Misumi Y., Ikehara Y.;
RA "Primary structure of rat liver dipeptidyl peptidase IV deduced from
RT its cDNA and identification of the NH2-terminal signal sequence as the
RT membrane-anchoring domain.";
RT J. Biol. Chem. 264:3596-3601(1989).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=88068516; PubMed=3479775;
RX Hong W., Doyle D.;
RA "cDNA cloning for a bile canalicular domain-specific membrane
RT glycoprotein of rat hepatocytes.";
RT Proc. Natl. Acad. Sci. U.S.A. 84:7962-7966(1987).
RN [3]
RN NUCLEOTIDE SEQUENCE OF 1-40.
RP MEDLINE=89034185; PubMed=3182821;
RX Hong W.J., Doyle D.;
RA "Membrane orientation of rat gp110 as studied by in vitro
RT translation.";
RT J. Biol. Chem. 263:16892-16898(1988).
RN [4]
RN PROTEIN SEQUENCE OF 28-58, AND TISSUE SPECIFICITY.
RX MEDLINE=90228896; PubMed=1970322;
RA McCaughan G.W., Wicks J.E., Creswick P.F., Gorrell M.D.;
RT "Identification of the bile canalicular cell surface molecule GP110 as
RT the ectopeptidase dipeptidyl peptidase IV: an analysis by tissue
RT distribution, purification and N-terminal amino acid sequence.";
RT Hepatology 11:534-544(1990).
RN [5]
RN PROTEIN SEQUENCE OF 281-302, AND MUTAGENESIS OF GLY-629; TRP-630;
RP SER-631; TRY-632 AND GLY-633.
RC TISSUE=Kidney;
RX MEDLINE=94128239; PubMed=7905271;
RA Iwaki-Egawa S., Watanabe Y., Fujimoto Y.;
RT "N-terminal amino acid sequence of the 60-kDa protein of rat kidney
RT dipeptidyl peptidase IV.";
RT Biol. Chem. Hoppe-Seyler 374:973-975(1993).
RN [6]
RN PROTEIN SEQUENCE OF 624-648.
RX MEDLINE=92190188; PubMed=1347701;
RA Ogata S., Misumi Y., Tsuji E., Takami N., Oda K., Ikehara Y.;
RT "Identification of the active site residues in dipeptidyl peptidase IV
RT by affinity labeling and site-directed mutagenesis.";
RT Biochemistry 31:2582-2587(1992).
RN [7]
RN SIGNAL-ANCHOR.
RP MEDLINE=90338089; PubMed=1974258; DOI=10.1083/jcb.111.2.323;
RX Hong W., Doyle D.;
RA "Molecular dissection of the NH2-terminal signal/anchor sequence of
RT rat dipeptidyl peptidase IV.";
RT J. Cell Biol. 111:323-328(1990).
CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from
CC polypeptides having unsubstituted N-termini provided that the
CC penultimate residue is proline.
CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide. Xaa-Yaa-|-

```
FT CONFLICT 54 54 Missing (in Ref. 4).
FT CONFLICT 183 183
FT CONFLICT 332 332 I -> T (in Ref. 2).
FT CONFLICT 332 332 T -> N (in Ref. 2).
FT CONFLICT 352 352 C -> V (in Ref. 2).
FT CONFLICT 394 394 V -> D (in Ref. 2).
FT CONFLICT 562 562 L -> F (in Ref. 2).
FT CONFLICT 624 624 R -> Q (in Ref. 2).
SQ SEQUENCE 767 AA; 88089 MW; ED947174F1F3E440 CRC64;

Query Match 86.5%; Score 3354.5; DB 1; Length 767;
Best Local Similarity 84.7%; Pred. No. 1.9e-220;
Matches 608; Conservative 52; Mismatches 55; Indels 3; Gaps 1;

Qy 1 NTPYKLYSRWISDHEHYLYKQENILVFNAYGNSVFLNSFTDFPGHSHINDYSIPD 60
Db 49 NTRFKVSYSLRWSDSEYLYKQENILVFNAYGNSVFLNSFTDFPGHSHINDYSIPD 108
Qy 61 GQFTLLYVYVQKWHYSYASYDYDLNKRQLITEERIPNNTQVWTSVGHKLYAYWNN 120
Db 109 RLFLVLLYVYVQKWHYSYASYDYDLNKRQLITEERIPNNTQVWTSVGHKLYAYWNN 168
Qy 121 DIYVKIEPNLPSVRAITWGKEDIYNGITDWYVEEVFSYALWSPNGTFFLAYAFND 180
Db 169 DIYVKIEPNLPSVRAITWGKEDIYNGITDWYVEEVFSYALWSPNGTFFLAYAFND 228
Qy 181 TEVPLIEYSFYSDESLOYPKTVRPYPKAGAVNPTVKFVVNTDLSSTVNTATSIQITAP 240
Db 229 TGVPLEIYSFYSDESLOYPKTVRPYPKAGAVNPTVKFVVNTDLSSTVNTATSIQITAP 288
Qy 241 ASMLIGHYICLVDTWATQERISLOWLRRIQNSYWDICDYDESSGRNCLVACQHEMST 300
Db 289 ASVTGTHYICLVDTWATQERISLOWLRRIQNSYWDICDYDESSGRNCLVACQHEMST 348
Qy 301 TGVWGRPRPSEPHFTLDGNSFYKIISNEEGRHICVFOIDKDO---CTEITKGTWEVIGI 357
Db 349 TGVWGRPRPSEPHFTLDGNSFYKIISNEEGRHICVFOIDKDO---CTEITKGTWEVIGI 408
Qy 358 EALTSDYLYYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELAPERCOYYSVSFSKEAY 417
Db 409 EALTSDYLYYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELAPERCOYYSVSFSKEAY 468
Qy 418 YQRCSPGLPLTYLTHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFILNETKFWYQ 477
Db 469 YQRCSPGLPLTYLTHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFILNETKFWYQ 528
Qy 478 MILPPHFDKSKYPLLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDCRGSYGQ 537
Db 529 MILPPHFDKSKYPLLDVYAGCSOKADTVFRLNWTYLASTENIIVASFDCRGSYGQ 588
Qy 538 DKIMHAINRLRGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSMLGSGGVF 597
Db 589 DKIMHAINRLRGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSMLGSGGVF 648
Qy 598 KGIAPVAPSRWEYSDVYTERYVGLPTPDNLDHYNSTVMSRAENFKQVYLLIHGTA 657
Db 649 KGIAPVAPSRWEYSDVYTERYVGLPTPDNLDHYNSTVMSRAENFKQVYLLIHGTA 708
Qy 658 DDNVHFQQAISKALVDGVDFQAMWYTDHEDGIASSTAHOHYTHMSHFICQCSL 715
Db 709 DDNVHFQQAISKALVDGVDFQAMWYTDHEDGIASSTAHOHYTHMSHFICQCSL 766
```

RESULT 8

```
ID_DP4_MOUSE STANDARD; PRT; 760 AA.
AC P28843;
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE IV) (T-cell activation antigen CD26) (thymocyte-activating molecule
DE (THAM) [contains: Dipeptidyl peptidase 4 membrane form (Dipeptidyl
DE peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form

DE (Dipeptidyl peptidase IV soluble form)]).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Swiss; TISSUE=Thymus;
RX MEDLINE=92129288; PubMed=1370813;
RA Marguet D.A., Bernard A.-M., Vivier I., Darmoul D., Naquet P.,
RA Pierres M.;
RT "CDNA cloning for mouse thymocyte-activating molecule. A
RT multifunctional ecto-dipeptidyl peptidase IV (CD26) included in a
RT subgroup of serine proteases.";
RL J. Biol. Chem. 267:2200-2208 (1992).
RN [2]
RP SEQUENCE REVISION.
RA Marguet D.A.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B10.A; TISSUE=Liver;
RX MEDLINE=95092780; PubMed=7999781;
RA Bernard A.-M., Mattei M.-G., Pierres M., Marguet D.;
RT "Structure of the mouse dipeptidyl peptidase IV (CD26) gene.";
RL Biochemistry 33:15204-15214 (1994).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldrelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kaspian A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bratt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltain L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshitoki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Heiton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Glickman J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalaka U., Smalilus D.B.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]

RP PROTEIN SEQUENCE OF 1-20.
RX MEDLINE-91302787; PubMed-1712807;
RA Vivier I., Marguet D.A., Naquet P., Bonicel J., Black D., Li C.X.-Y.,
RA Bernard A.-M., Gorvel J.-P., Pierres M.;
RA "Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl
RT peptidase IV).";
CC J. Immunol. 147:447-454(1991).

CC -I- FUNCTION: Removes N-terminal dipeptides sequentially from
CC polypeptides having unsubstituted N-termini provided that the
CC penultimate residue is proline.
CC -I- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither pro nor hydroxyproline.
CC -I- SUBUNIT: Homodimer.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC soluble form.
CC -I- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC by proteolytic processing.
CC -I- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC -----
DR EMBL; X58384; CAA41274.1; -; Genomic DNA.
DR EMBL; U12620; AAA82213.1; -; Genomic DNA.
DR EMBL; U12599; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12600; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12601; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12602; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12603; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12604; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12605; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12606; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12607; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12608; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12609; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12610; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12611; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12612; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12613; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12614; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12615; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12616; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12617; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12618; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12619; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; AK085370; BAC39434.1; -; mRNA.
DR EMBL; BC022183; AAH22183.1; -; mRNA.
DR HSSP; P27487; 1PFO.
DR SMR; P28843; 37-759.
DR MEROPS; S09.003; -.
DR Ensembl; ENSMUSG00000035000; Mus musculus.
DR MGI; MG1:94919; Dpp4.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0046581; C:intercellular canalliculus; IDA.
DR InterPro; IPR002471; Pept S9 AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.

DR InterPro; IPR000379; Ser.estr.
DR Pfam; PF00330; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;
KW Protease; Serine protease; Signal-anchor; Transmembrane.
FT CHAIN 1 760 Dipeptidyl peptidase 4 membrane form.
FT CHAIN 37 760 Dipeptidyl peptidase 4 soluble form (By
FT similarity).
FT TOPO_DOM 1 6 Cytoplasmic (Potential).
FT TRANSMEM 7 28 Signal-anchor for type II membrane
FT protein (Potential).
FT TOPO_DOM 29 760 Extracellular (Potential).
FT COMEBIAS 295 466 Cys-rich.
FT ACT_SITE 624 624 Charge relay system (By similarity).
FT ACT_SITE 702 702 Charge relay system (By similarity).
FT ACT_SITE 734 734 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 83 83 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 90 90 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 213 213 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 223 223 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 315 315 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 328 328 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 514 514 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 679 679 N-linked (GlcNAc...) (By similarity).
FT DISULFID 379 388 By similarity.
FT DISULFID 438 441 By similarity.
FT DISULFID 448 466 By similarity.
FT DISULFID 643 756 By similarity.
SQ SEQUENCE 760 AA; 87437 MW; ASF644B46B43DF8 CRC64;
Query Match 86.1%; Score 3340; DB 1; Length 760;
Best Local Similarity 84.9%; Pred. No. 1.8e-219;
Matches 607; Conservative 55; Mismatches 49; Indels 4; Gaps 1;
QY 1 NTRYRLKSLRWISDHEYLKQENNLVFNABYGNSSVFLENSTDFEGHSINDYSISPD 60
DB 49 STFRVKSLSLWVSDFEYLKQENLLLNLAHGNSSIFLENSTFESFGY----HSVSPD 104
QY 61 GQFILLENNYKQWRHSYASDYIDLNKRQLITEIRIPNNQWYTSVGHKLAYVWN 120
DB 105 RLFLVLENNYKQWRHSYASVNIYDVKRQLITEKIPNNQWYTSVGHKLAYVWN 164
QY 121 DIIYKLENNLPSYRITWTCKEDIYNGITDWWYEEVSAYSALWNSNGTFLAYAQFND 180
DB 165 DIIYKVEPPLPSHRITSTGEENVYNGITDWWYEEVFGAYSALWNSPNTFLAYAQFND 224
QY 181 TSPVLIYSFYSDLSQVPTKVRVPYKAGAVNPTVKPVVNTDSLSSVTNATSIQITAP 240
DB 225 TGVPLIYSFYSDLSQVPTKVTWIPYKAGAVNPTVKPVVNTDSLSSSSSAPIQIPAP 284
QY 241 ASMLIGDHYLCDVWTWATERISLOWLRRIQNYSVMDICDYDESSGRWNCVLAQRIEMST 300
DB 285 ASVARGDHYLCDVWATERISLOWLRRIQNYSVMAICDYDKINLTWNCPSQQQVEMST 344
QY 301 TGWGRFRPSEPHFTLDGNSFYKISNEEYRHICYFQIDKDKCTFITKGTWEVIGIAL 360
DB 345 TGWGRFRPAEPHFTSDGSSFYKISDKDGYKHCHFPKDKDCTFITKGAWEVISIAL 404
QY 361 TSDYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCENLPERCOVYSFSEKAYYQL 420
DB 405 TSDYLYISNQYKPNPGRNLYKIQLDHTNVKCLSCDLPNPERCOYAVSFSEKAYYQL 464
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWQML 480
DB 465 GCMGPGPLPLYTLHRTDHLKRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWQML 524
QY 481 PPHFDKSKYPLLLDVIYAGPSQKADTVFRLNWAYTYLASTENIIVASPDGRSGVQGDKI 540
DB 525 PPHFDKSKYPLLLDVIYAGPSQKADAFRLNWAYTYLASTENIIVASPDGRSGVQGDKI 584
QY 541 MHAINRRLLGTPEVEDQIEAARQFSXMGFVDNKRITAIWGSYGYVTSVWLGSVGSVFKCG 600

Db 585 MHAINRLGLTLEVEDQIEAARQFVKMGFVDSKRVIAIWGWSYGGYVTSNVLGSGSVFKCG 644
Qy 601 IAVAPSRWEYDYSVYTERYMGTPEDNLDHYRNSVTMSRAENFKQVYELLHGTADDN 660
Db 645 IAVAPSRWEYDYSVYTERYMGTPEDNLDHYRNSVTMSRAENFKQVYELLHGTADDN 704
Qy 661 VHQFQSAQISKALVDGVDFQAMWYTDHGHASSTAHOIYTHMSHFIIKQCFSL 715
Db 705 VHQFQSAQISKALVDGVDFQAMWYTDHGHASSTAHOIYTHMSHFIIKQCFSL 759

RESULT 9
Q75S83_AKXHB
ID Q75S83_AKXHB PRELIMINARY; PRT; 751 AA.
AC Q75S83;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Dipterydylpeptidase 4a (EC 3.4.14.5).
GN Name=DPP4a;
OS Agkistrodon haly's brevicaudus (Korean slamosa snake) (Gloydus haly's brevicaudus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Gloydus.
OX NCBI_TaxID=259325;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Venom gland;
RA Ogawa Y., Yanoshita R., Murayama N., Higuchi S., Samejima Y.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB158224; BAD06332.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004287; F:hydrolase activity; IEA.
DR GO; GO:0016787; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 751 AA; 86127 MW; 6EA3F2D829BE97D4 CRC64;

Query Match 64.2%; Score 2490.5; DB 2; Length 751;
Best Local Similarity 62.9%; Pred. No. 2e-161;
Matches 449; Conservative 110; Mismatches 144; Indels 11; Gaps 5;

Qy 3 YRLKLSLWISDHEVLYKQENILVFNABYGNSSVPLENSTDFGHSINDYSISPDGQ 62
Db 49 FOYKSNLWMSGHEVYTNQNNVLLYNDRESIVLSNDTLDSEFSS--QAILSPDRK 106

Qy 63 FILLEYNVKWRHSYTA5DIYDLAKRQLITERRIPNNTQWTSVPVGHKLAVVWVNDI 122
Db 107 FALQYSYKWRHSYTA5THIYDLNRTKITENPLPTNIQYISWSVPVGHKLAVVYRNV 166

Qy 123 YVKIEPNLPSYRITGTGKEDIIYNGITDWWYEEVFSAYSGALWMSNGTFLAYAQNDTE 182
Db 167 YVKATPNASPVQITENGAEKNILGLADWYEEEMFGTHSALWSPNGRFLAEINDTE 226

Qy 183 VPLIEVSFYDESLOYPKTVRVPYKAGAVNPTVKFVNVNTDSSLVSNTAITSQITAPAS 242
Db 227 VPWMEYSFYSEDLTQYPKTIKIPYKAGAINPTIRFLVLDI----SLSPKNISEIVAPSS 282

Qy 243 MLIGDHYLCVWTATQERISLOWLRIONYSVMDICDYDESSGRWNCVLAQHIENSTTC 302
Db 283 IISGDHYLSAVTWVTERICQWLRRIQNFSLVITICDY---SGAWHCPKREHLEESKT 339

Qy 303 WVGFRFPSEPHFTLDGNSFYKLIISNBEGYRHCYFQIDKDKDCTFITKGTWETVIGIEALT 362

Db 340 WVGFRFPSEPYFTSDKISYRYRIISDSGYKHIHYTDSAGK-VKPIITSGKWEVISISAVTN 398
Qy 363 DYLYYSINSEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSEAKYQJLRC 422
Db 399 NSLYFISNEPEGRPGGRHLYKVDLKNLKKICITCNSEKACQYFSVSFSDSRYKLNLC 458
Qy 423 SGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQPSKKLDFIILNETKFWYQMLPP 482
Db 459 YGDLPYFTLQNSITDKAIKTLEDNNLNKVLKEIQMPCRLSNITLHGQTYWYQMLPP 518
Qy 483 HFDKSKYKPLLLDVYAGPCSQKADTVFRLNWAYTLASTENIIVASFDGRSGYQGDIMH 542
Db 519 NFDESXKYPLLDVYAGPCSQKADAAFRINWSTYLASSEGLIVASFDGRSGYQGDIMH 578
Qy 543 AINRRLGTPEVEDQIEAARQFVKMGFVDSKRVIAIWGWSYGGYVTSNVLGSGSVFKCGIA 602
Db 579 AIYRRLGTPEVEDQIEAARQFVKMGFVDSKRVIAIWGWSYGGYVTSNVLGSGSVFKCGIA 638
Qy 603 VAPSRWEYDYSVYTERYMGTPEDNLDHYRNSVTMSRAENFKQVYELLHGTADDNVH 662
Db 639 VAPSRWEYDYSVYTERYMGTPEDNLDHYRNSVTMSRAENFKQVYELLHGTADDNVH 698
Qy 663 FQFSAQISKALVDGVDFQAMWYTDHGHASSTAHOIYTHMSHFIIKQCFSLP 716
Db 699 FQFSAQISKALVDGVDFQAMWYTDHGHASSTAHOIYTHMSHFIIKQCFSLP 751

RESULT 10
Q75S82_AKXHB
ID Q75S82_AKXHB PRELIMINARY; PRT; 751 AA.
AC Q75S82;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Dipterydylpeptidase 4b (EC 3.4.14.5).
GN Name=DPP4b;
OS Agkistrodon haly's brevicaudus (Korean slamosa snake) (Gloydus haly's brevicaudus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Gloydus.
OX NCBI_TaxID=259325;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Venom gland;
RA Ogawa Y., Yanoshita R., Murayama N., Higuchi S., Samejima Y.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB158225; BAD06333.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004287; F:hydrolase activity; IEA.
DR GO; GO:0016787; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 751 AA; 86155 MW; DAF219B4FEE7629A CRC64;

Query Match 64.2%; Score 2489.5; DB 2; Length 751;
Best Local Similarity 62.9%; Pred. No. 2.4e-161;
Matches 449; Conservative 110; Mismatches 144; Indels 11; Gaps 5;

Qy 3 YRLKLSLWISDHEVLYKQENILVFNABYGNSSVPLENSTDFGHSINDYSISPDGQ 62
Db 49 FOYKSNLWMSGHEVYTNQNNVLLYNDRESIVLSNDTLDSEFSS--QAILSPDRK 106

QY 63 FILLEYVYKQWRHSYTSYDIYDLNKRQLITEIRPNNTOWTWVSPVGHKLAAYVNNDI 122
DB 107 FALQVSYKVRHRSYSTASHIYDLNRRTKITENPLPTNIQIYSWSVPGHKLAYVRRNV 166
QY 123 YKIEBNLPYSRITWTGKEDIIYNGITDWYEEVEFSAYSALWSPNGTFLAYAOFNDE 182
DB 167 YKATNPNASPVQITENGAKNKLGLADWYEEEMFGTHSALWSPNGRFLAFENDE 226
QY 183 VPIEYFSYDSBLSQVKTVPVYPKAGAVNPVFPVNTDLSLSTVNTATSIQITAPAS 242
DB 227 VPVMEYSFSEDTLQYPKTIKIPYKAGAINPTIRLFVLDI---SLSPKNTSEIVAPSS 282
QY 243 MLIGDHYLVDVWATQERISLOWLRIONYSVMDICDYDESSGRWNCVLAHQHIESTTG 302
DB 283 IISGDHYLSVVTWVTDERICVQWLRRIQNPFSVLITCIDY---SGAWCPKEREHLESKTG 339
QY 303 WGRFRPSPHPFTLDGNSFYKIISNEEGYRHICYFOIDKKDOCTFITKGTWEVIGIBALTS 362
DB 340 WGRFQPSPEYFTSDKISYRIISDSEGYKHITDTSAGK-VKPIITSGKWEVISISAVTN 398
QY 363 DYIYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYVSFSKBAKYQLRC 422
DB 399 NSLYFISNPEGRPGGRHLYKVDLKNLKKICITCNSKEAACQYFSVSPSTDSRYKKLC 458
QY 423 SGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVOMPSKKLDFILNETKFWOMILPP 482
DB 459 YGDLPLYTLQNSITDOKAIKTELDNNLNKVLKBIQMPCKRLSNITLHGQTYWYQMLPP 518
QY 483 HFDKSKYPLLLDVYAGPQSOKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKIMH 542
DB 519 NFDESKYPLLDVYAGPQSOKADAFRINWSTYLASSSEGIIVASFDGRGSGYQGDKILH 578
QY 543 AINRRLGTVEVEDQISAAKQFSKMGFVNDKRIAWGWSYGGYVTSVWLGSVFKCGIA 602
DB 579 AIYRRLGTVEVEDQISAAKLFSEMSFVDKRIAIGWWSYGGYVTSVWLGSVFKCGIA 638
QY 603 VAPVSRWEYDYSVYTERYVGLTPEDNLDHNRSTVMSRAENPKQVEYLLIHGTADDNVH 662
DB 639 VAPVSRWQYDYSIYTERYVGLTPEDNLDHNRSTVMSRAENPKQVEYLLIHGTADDNVH 698
QY 663 FQQAQISKALVDVGVDFQAMWYTDHEDHGIASTAHQIYTHMSHPFKOCFSLP 716
DB 699 FQQAQISKALVDVGVDFQAMWYTDKDHGIGGH-AHSHIYQHSHPFKOCFKLP 751

RESULT 11

Q52181 CHICK
ID Q52181_CHICK PRELIMINARY; PRT; 759 AA.
AC Q52181_ CHICK PRELIMINARY; PRT; 759 AA.
DT 25-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04.29g21;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.,
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT gene function analysis".
RL Genome Biol. 6:R6-R6(2005).
DR EMBL; AJ270903; CAG32562.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_eatra.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; I.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Hypothetical protein; Transmembrane.
SQ SEQUENCE 759 AA; 86343 MW; B9A0C38CCB8644AC CRC64;

Query Match

Best Local Similarity 63.9%; Score 2478; DB 2; Length 759;

Matches 466; Conservative 92; Mismatches 147; Indels 16; Gaps 9;

QY 1 NTRYLKLRLSRWISDHEVLYKQEN-NILVFNAEYGNSSVFLNENSTPDERGHSINDYSISP 59
DB 50 NDYVYKTNLQWISGNQVILHETSNGNLLRFDAETGTSSVLLNTTISI--HEATTAILSP 107
QY 60 DQGFILLEYVYKQWRHSYTSYDIYDLNKRQLITEIRPNNTOWTWVSPVGHKLAAYVNN 119
DB 108 DQRFALQVYKYEKLWRHSYTSYDIYDFNTSSILDALLPNDTQVISWSPVGHKLAAYVNN 167
QY 120 NDIYVYKIEPNLPSYRITWTGKEDIIYNGITDWYEEVEFSAYSALWWSPNGTFLAYAOEN 179
DB 168 NNIIYKASPTAAPVQITNSGENKIFNGIPDWYEEEMFGSHALWWSPNGNPFVAYAFEN 227
QY 180 DTEVPLIEYSFYSDLSQVPTVRVPYKAGAVNPVFPVNTDLSLSTVNTATSIQITAPAS 239
DB 228 DTEVPEVIEYSFSEDTLQYPKTIKIPYKAGAKNPTVKEFIVD---IQMLPDFNTEISP 284
QY 240 PASMILGHYLCVDTWATQERISLOWLRIONYSVMDICDYDESSGRWNCVLAHQH--I 296
DB 285 PAEIKSGDHYLSVVTWVTDERICLQWLRRIQNSVLTICDPESATGNMTC--HRKNPFWK 342
QY 297 ENSTTGWGRPRSPHPFTLDGNSFYKIISNEEGYRHICYFOIDKKDOCTFITKGTWEVIG 356
DB 343 KVQLAGN--QISAIPLFAPDNTTYIKVFSNTEGYKHIHYINGTEAPVP-ITEGRKEVIS 399
QY 357 IEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVT-CLSCELNPERCOYYVSFSKBA 415
DB 400 IAAVTKYFLYIISNQGEMPGGRNLYKMLLESSPKSTQCVSCDLNQERCOYYVSASFSDA 459
QY 416 KYIYQRCSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVOMPSKKLDFILNETKFW 475
DB 460 QYYQLNCLGPGLPMTSLHRSSDDQVLRVLENNTELENSLKDIQMPSKKLGSIITVGGYNLW 519
QY 476 YQMILPPHFDKSKYPLLLDVYAGPQSOKADTVFRLNWTATYLASTENIIVASFDGRGSGY 535
DB 520 YQMILPPHLDSSKKYPLLEVVYAGPQSOKQVDHVFRLNWTATYLASTEQIIVASFDGRGSGY 579
QY 536 QGDKIWHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAWGWSYGGYVTSVWLGSVSGS 595
DB 580 QGDEIWHAINRRLGTFEVEDQISAAFTSEMSFVDKRIAIGWWSYGGYVTSVWLGSVSGS 639
QY 596 VFKEGIAVAPVSRWEYDYSVYTERYVGLTPEDNLDHNRSTVMSRAENPKQVEYLLIHG 655
DB 640 VFKEGIAVAPVSRWQYDYSIYTERYVGLTPEDNLDHNRSTVMSRAENPKQVEYLLIHG 699
QY 656 TADDNVHFQOASQAIKALVDVGVDFQAMWYTDHEDHGIASTAHQIYTHMSHPFKOCFSL 715
DB 700 TADDNVHFQOASQAIKALVDVGVDFQAMWYTDKDHGIGGH--SQAHKHIIYTHMSHPFKOCFSL 758
QY 716 P 716
DB 759 P 759

RESULT 12

Q641D6 XENLA
ID Q641D6_XENLA PRELIMINARY; PRT; 737 AA.
AC Q641D6_

Db 462 YALVCYGPPISTLHGRDQEIKILEENKELENALKNLQLPKBEIKKLEVDITLWYK 521
Qy 478 MILPPHDKSKYPLLLDVYAGCSQKADTVFRLNMTATYLASTENIIIVASFDCRGSGYQG 537
Db 522 MILPPQFDRSKYPLLIQVYGGCSQSVRSVFAVNWISYLSKSGWIALVDGRGTAFQG 581
Qy 538 DKIMHAINRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYTSMLVSGSGGVF 597
Db 582 DKLYAYVRKLGVEVEDQITAVRKFTMGFIDFKRIAIWGSYGGVYSSALASGTLGF 641
Qy 598 KCGIAPVPSRWEYSDVSYTYRYMGLPTPDNLDHYRNSVMSRAENFKOVEYLLIHGTA 657
Db 642 KCGIAPVPSRWEYASYTYRYMGLPTPDNLDHYRNSVMSRAENFKOVEYLLIHGTA 701
Qy 658 DDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIHYTHMSHFIKQCFSL 715
Db 702 DDNVHFQOQAIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLTYTHMTHFLKQCFSL 758

RESULT 15
Q8R492 RAT PRELIMINARY; PRT; 761 AA.
AC Q8R492; 2006 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DE 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Fibroblast activation protein alpha subunit.
GN NamesPap;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zuckerswerdt K., Park J.E.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF493782; AAM11677.1; -; mRNA.
DR HSSP; P27487; 1PQ.
DR MEROPS; S09.007; -;
DR Ensembl; ENSRNOG00000005679; Rattus norvegicus.
DR RGD; 621253; Fap.
DR GO; GO:0016020; C.membrane; IEA.
DR GO; GO:0004177; F.aminopeptidase activity; IEA.
DR GO; GO:0004274; F.dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F.hydrolase activity; IEA.
DR GO; GO:0004287; F.prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P.proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept S9 AS.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_eatrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 761 AA; 87843 MW; EB214BBB4BB14F2C CRC64;

Query Match 55.2%; Score 2141.5; DB 2; Length 761;
Best Local Similarity 54.0%; Pred. No. 1.5e-137;
Matches 388; Conservative 128; Mismatches 189; Indels 13; Gaps 6;

Qy 2 TYRLKLYSLRWISDHEVLYKQEB-NNIIAVFNAEYGNSSVFLNSTPDEFHGSIN--DYSIS 58
Db 51 TFSYKTYFPNWISQEYVHQSEDDNIVFYNIETRESYIILSNSTM----KSVNATDGLS 106
Qy 59 PDQQFILLEYNVYKQWRHSYASYDIYDLNKRQLITEERIPNNNTQWTVSPVGHKLAYVW 118
Db 107 PDRQFYLESDYSKLWRYSTATYIYDLQNGEFVRGYELPRPIQYLWSPVGSKLAYV 166
Qy 119 NNDIYKIEPNLPSRITWTGKEDIYNGITDWWYEEVFSAYSAWSPNGTFLAYAO 178
Db 167 QNNIYLKQRPDPFPQITTYGRENRIENGIPDWWYEEEMLAATKYALWSPDGKLYAYVF 226

Qy 179 NDETEVLIBSYFSDSLSQVPTKTRVPYPKAGAVNPTVKFFVVTNDSLSVNTATSIQIT 238
Db 227 NDSIDIPIIAYSYYGDG--QYPRITINPYPKAGAKNPIVRVFIVDITYPHHV---GRIEVP 281
Qy 239 APASMLIGHYLCDVTWATQERISLOWLRRIQNYVMDICDYDESSGRWNCLVAROHIE 298
Db 282 VPEMIASDSYFTWLTWVTNERVCLQWLKRVQNVSVLSICDPREDHAMDCPKNQBHIE 341
Qy 299 STTGWVGRFRPSPHFTLDGNSFYKLIISNEEGVYRHCYQIDKKDCTFFITKGTWEVIG 358
Db 342 SRTGWAGGFVSTPAFSQDAASYIKFSDKDGKYKHIIKDTVENAIQITSGKWEALYIF 401
Qy 359 ALTSYLYIISNEYKMGPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKY 417
Db 402 RVTQDSLFYSSNEFEGYPGRRNIYRISIGNSPSKCVTCHLRKERCQYYTASFYKAKY 461
Qy 418 YQURCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFILNETFWYQ 477
Db 462 YALICYGPGLPISLTHDGRDQEIQVLEENKELENALRNLIQLPVAVIKKLEDGGMTFWYK 521
Qy 478 MILPPHFDKSKYPLLLDVYAGCSQKADTVFRLNMTATYLASTENIIIVASFDCRGSGYQG 537
Db 522 MILPPQFDRSKYPLLIQVYGGCSQSVRSVFAVNWISYLSKSGWIALVDGRGTAFQG 581
Qy 538 DKIMHAINRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYTSMLVSGSGGVF 597
Db 582 DKFLHAVYRKLGVYEVEDQITAVRKFTMGFIDEGRIALWGSYGGVYSSALASGTLGF 641
Qy 598 KCGIAPVPSRWEYSDVSYTYRYMGLPTPDNLDHYRNSVMSRAENFKOVEYLLIHGTA 657
Db 642 KCGIAPVPSRWEYASYTYRYMGLPTPDNLDHYRNSVMSRAENFKOVEYLLIHGTA 701
Qy 658 DDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIHYTHMSHFIKQCFSL 715
Db 702 DDNVHFQOQAIAKALVNAQVDFQAMWYSDQNHGISGRSQNHLYTHMTHFLKQCFSL 759

Search completed: January 30, 2006, 14:20:09
Job time : 167 secs

THIS PAGE BLANK (USPTO)